This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2025.3604710

LOGO

Revolutionizing Wearable Sensor Data Analysis with an Automated Decision-Making Model for Enhanced Human Activity Detection

Nitesh Bharot, *Sr. Member, IEEE*, Priyanka Verma, *Sr. Member, IEEE*, Ankit Vidyarthi, Deepak Gupta *Sr. Member, IEEE, John G. Breslin, Sr. Member, IEEE*

***Abstract*— Human Activity Recognition (HAR) stands as a crucial technology, with applications ranging from healthcare monitoring to sports analytics. However, the traditional approach to HAR is often time-consuming and susceptible to human errors due to the high complexities involved in processing diverse sensor data. Recognizing the imperative for efficiency and accuracy in HAR systems, we propose the development of an Automated Decision-maker (ADM) system. This system serves to automate HAR pipelines, addressing the challenges posed by the huge sensor data. By harnessing the power of automation, ADM significantly streamlines the HAR process, reducing the time required for hyperparameter tuning and minimizing the risk of human errors. The results obtained from our proposed ADM system demonstrate notable improvements in HAR performance, showcasing achieved accuracy of 96.436% for UCI-HAR & 99.783% for PAMAP2 datasets. Moreover, ADM can be described as an innovative approach that contributes to the optimization of HAR systems while also establishing a foundation for building robust and reliable systems in complex environments.**

***Index Terms*— Healthcare, UCI-HAR, PAMAP2, AutoML, activity recognition, IoT, sensors**

I. Introduction

In the evolving landscape of computation, Human Activity Recognition (HAR) [1], [2] is emerging as a cornerstone technology with widespread applications across healthcare [3], sports [4], industry, and human-computer interaction [5], [6]. By leveraging advancements in the Internet of Things (IoT) and Artificial Intelligence (AI), HAR systems aim to interpret human activities through data captured by an array of sensors. These activities range from routine daily actions to complex behaviors in specialized environments, underpinning innovations that promise to transform safety measures, healthcare monitoring, sports analytics, and interactive technologies.

At the heart of HAR's progress are sensors—both wearable and environmental that continuously collect data on human movements and behaviors. Wearable sensors, including Inertial Measurement Units (IMUs) [7], [8], [9], offer a portable and unobtrusive means to gather activity data, while environmental sensors, such as cameras and millimeter radar sensors, provide contextual insights into human actions. As the integration of sensor technologies becomes increasingly prevalent in healthcare applications, there arises a pressing need for an automated system to streamline the complex pipelines involved in HAR systems. This paper introduces an innovative solution, Automated Decision Maker (ADM), aimed at automating solution pipelines for the HAR system. By aiming to resolve the challenges with sensor data & human involvement, ADM seeks to revolutionize HAR methodologies, providing a more efficient, accurate, and robust approach for real-world applications in healthcare and beyond. This approach not only promises to enhance the accuracy of activity recognition but also aims to build more robust and stable HAR systems capable of adapting to the dynamic nature of human behaviors [10].

The main contributions of this work are:

- The novel approach introduced eliminates the need for intricate hyperparameter tuning, streamlining the model development process. By automating this crucial aspect, the proposed methodology minimizes the potential for human errors in the hyperparameter selection phase, enhancing the reproducibility and reliability of the HAR system.
- It focuses on the development of a novel model that not only enhances accuracy but also excels in other metrics on UCI-HAR and PAMAP2 datasets.
- The elimination of manual hyperparameter tuning represents a significant advancement, making the proposed HAR approach more robust, less prone to subjective biases, and ultimately contributing to the broader goal of creating reliable and automated HAR systems.

The remaining paper is organized as section 2 presents the related work and section 3 presents the proposed methodology. Section 4 shows the results achieved and section 5 concludes

This publication is supported by SFI under Grant 12/RC/2289-P2 (Insight).

Nitesh Bharot and John G. Breslin is with Data Science Institute, University of Galway, H91TK33, Ireland (e-mail: firstname.lastname@universityofgalway.ie)

Priyanka Verma is with School of Computer Science, University of Galway, H91TK33, Ireland (e-mail: priyanka.verma@universityofgalway.ie)

Ankit Vidyarthi is with Department of CSE & IT, Jaypee Institute of Information Technology, U.P. India (e-mail: dr.ankit.vidyarthi@gmail.com)

Deepak Gupta is with Department of CSE, Maharaja Agrasen Institute of Technology, India (e-mail: drdeepakgupta.cse@gmail.com)

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2025.3604710

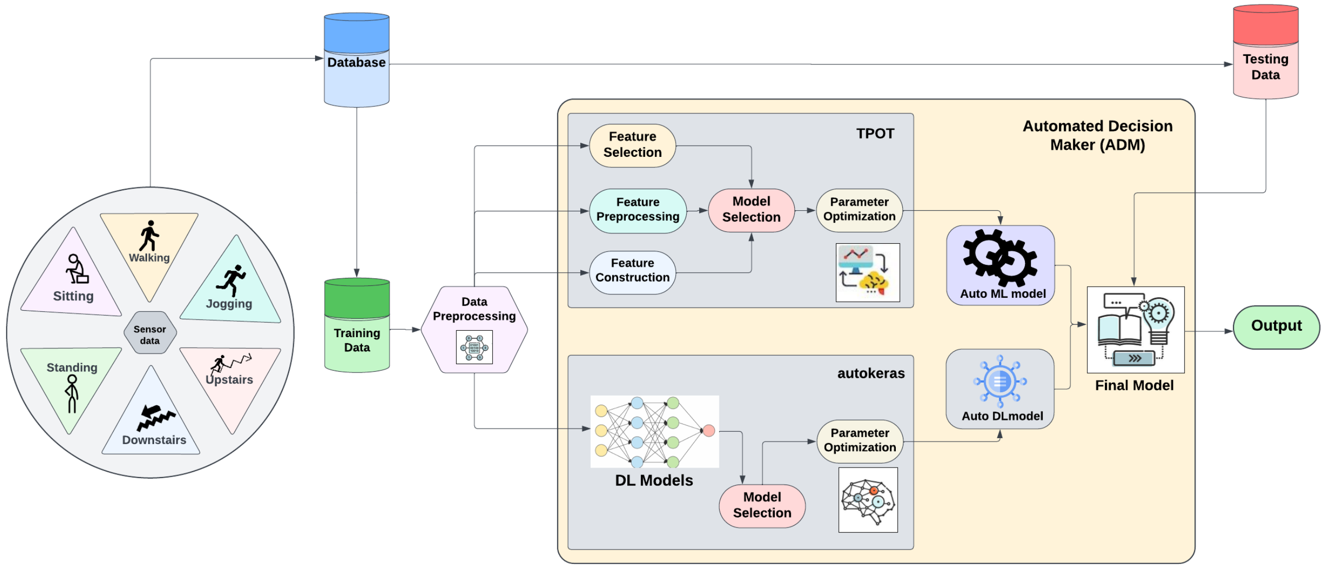

Fig. 1. Proposed methodology for human activity detection

the work.

II. Related Work

In the realm of HAR, an array of Machine Learning (ML) models have been employed, each harnessing various sensing technologies to discern and classify human behaviors [11]. While camera-based HAR has garnered considerable attention, recent research has pivoted towards embedded sensors, propelled by the challenges posed by high computational resource requirements and burgeoning concerns about user privacy [12], [13].

In the pursuit of understanding human behavior, linear models, such as logistic regression, have been harnessed for multinomial classification in HAR, as demonstrated by Fan et al. [14]. This technique, categorized under linear model classification, utilizes the LIBLINEAR optimizer to fine-tune model parameters, showcasing the adaptability of linear models in capturing and classifying diverse human activities.

Ensemble classifiers have emerged as a powerful tool in HAR, capitalizing on the strengths of multiple models. The ensemble approach combines specialized models, each excelling in discerning specific instances, leading to an overall improved model. Tan et al. [15] presents an ensemble of gated recurrent unit (GRU), a convolutional neural network (CNN) and Deep Neural Network (DNN) for activity recognition from smartphone sensor data.

Deep Learning (DL) is emerged as a transformative paradigm in HAR systems, demonstrating its prowess in learning intricate features directly from raw sensor data and achieving state-of-the-art performance. Researchers have explored diverse DL architectures, each contributing to the advancement of HAR methodologies [16], [17], [18], [19], [20].

Khan et. al. [21] introduced an attention-based multi-head Convolutional Neural Network (CNN) model tailored for HAR. This innovative approach selectively focuses on essential information, enhancing recognition performance using significant elements of the input data. Meanwhile, Cruciani et. al. [22] delved into the use of pre-trained CNN models for large-scale HAR datasets, harnessing knowledge gained from extensive data to capture discriminative features effectively.

Wan et. al. [23] proposed a CNN-based architecture prioritizing extraction of local features. Recognizing importance of spatial information, this approach adeptly classifies human activities with accuracy. Furthermore, Mutegeki et. al. [24] and Deep et. al. [25] embraced CNN-Long Short-Term Memory (LSTM) architectures, combining spatial and temporal information for superior activity prediction. These models excelled by capturing both short and long-term dependencies in the input data.

On the other front, the integration of DL algorithms into ensemble models has garnered attention for enhancing robustness and performance in ML applications, including HAR [16], [17], [18], [19], [20]. While ensemble models effectively leverage the learning capabilities of multiple learners, they introduce computational challenges and often adhere to conventional approaches in learner selection.

While previous studies have explored various ML and DL models, including logistic regression, CNNs, LSTMs, and ensemble approaches for HAR using wearable sensors, most still require manual intervention in hyperparameter tuning, model selection, and feature engineering. In contrast, our proposed Automated Decision Maker (ADM) integrates AutoML (TPOT) and AutoDL (AutoKeras) to fully automate the pipeline. This not only improves performance but also minimises human bias and enhances reproducibility.

III. Proposed Methodology

This section delves into the proposed methodology, describing data preprocessing, automated decision-maker, and workflow of the proposed methodology. Here Algorithm 1 describes the proposed methodology and Fig. 1 presents the flow and diagrammatic overview of proposed approach.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2025.3604710

Algorithm 1 Proposed Methodology

Input: Training data 'T' & testing data 'τ'
Output: Trained Automated Decision Maker (ADM)

1: Collect data D from sensors
2: Use train_test_split to divide the data into training & testing database
    $T, \tau = train\_test\_split(D)$
    $T_X, T_y \ = \ T.iloc[:, : -1], \ T.iloc[:, -1]$
3: Use mean imputation, standard scaling and isolation forest for outlier detection
    T_X = Mean_imput(T_X)
    T_X = Standard_scaler().fit_transform(T_X)
    is = Outlier_detection(T_X, T_y)
    indexes = Outlier_indices
    $T_X \ = \ T_X.drop(indexes, axis = 0)$
    $T_y \ = \ T_y.drop(indexes, axis = 0)$
    T = pd.concat([T_X, T_y],axis = 1)
    T = T.resample(frac=1)
    T.reset_index(drop=True,inplace=True)
    T_X, T_y = T.iloc[:,-1],T.iloc[:,-1]
4: Pass the data to ADM (Λ)
    cv = Repeated_Stratified_K_Fold(n_splits=10, n_repeats = 3,random_state=1)
    Φ = TPOTClassifier(generations=10, population_size=100,cv= cv,scoring=accuracy)
    ϕ = StructuredDataClassifier(overwrite=True, max_trials=15).fit(T_X, T_y,epochs=10, validation_split=0.15)
    $\tau_X, \tau_y \ = \ \tau.iloc[:, : -1], \tau.iloc[:, -1]$
    $\Lambda \ = \ best\_of(\Phi, \phi, \tau_X, \tau_y)$
5: return Λ

Initially, the data, 'D', collected from the sensor is bifurcated to form the training and testing dataset. This training dataset 'T' is then further passed through the preprocessing phase to provide insights into data and then this preprocessed data is passed to the ADM for training and validation purposes. The distinguished feature of this ADM lies in the fact that it eliminates the need for feature selection, feature preprocessing, and model selection that was previously done by humans. This reduces human error and the novel ensemble method of autoML and AutoKeras ensures that all the ML and DL models are taken into picture for finalizing the best-suited model.

A. Data Preprocessing Phase

This phase deals with the employment of a comprehensive data preprocessing strategy for enhancing the efficacy of the classification model. This encompasses mean imputation for handling missing values and standard scaler normalization to address varying feature scales and the application of isolation forests for outlier detection:

$$T_{X_{ij}} = \frac{T_{X_{ij}} - \bar{T_{X_i}}}{\sigma_{T_{X_i}}} \tag{1}$$

where $T_{X_{ij}}$ is the j^{th} element of the i^{th} column, i is the column number, j represents the index number, $\bar{T_{X_i}}$ represents the mean of values of the i^{th} column, to address varying feature scales and the application of isolation forests for outlier detection.

The Isolation Forest (iForest) algorithm efficiently isolates outliers by leveraging their inherent characteristics: being few in number and significantly different. It operates in two stages: first, constructing isolation trees (iTrees) using a subsample of the training data, and second, testing instances through these iTrees to compute outlier or anomaly scores. The algorithm's efficiency in isolating anomalies makes it highly effective for outlier detection. Additionally, its linear time complexity and low memory usage make it ideal for handling large-scale datasets.

During training, iTrees are built by recursively partitioning the data based on random features until instances are isolated or a predefined tree height is reached, creating partial models.

$$Height\_limit = ceil(log_2(sample\_size)) \tag{2}$$

The outlier score of the data point is based on the structure of the iTrees which resembles the structure of the Binary Search Trees (BST). Therefore, the estimation of the anomaly/outlier score is given as:

$$o(T_X, s) = 2^{-E(a(T_X))/a(\bar{T}_X)} \tag{3}$$

where T_X is the instance to be tested, is the average of $a(T_X)$ from a collection of iTrees, $E(a(T_X))$ $a(\bar{T}_X)$ represents the average of $a(T_X)$ given n number of instances, and $a(T_X)$ is the path length of an instance T_X. Here

$$a(\bar{T}_X) = 2 * Hnumber(n-1) - (2(n-1)/n) \tag{4}$$

$$Hnumber(x) = log_e(x) + \varepsilon \tag{5}$$

Where Euler's constant, ε = 0.5772156649. General values of Eq.3 are:

$$E(a(T_X)) \rightarrow a(\bar{T}_X), o(T_X, s) \rightarrow 0.5 \tag{6}$$

$$E(a(T_X)) \rightarrow 0, o(T_X, s) \rightarrow 1 \tag{7}$$

$$E(a(T_X)) \rightarrow n-1, o(T_X, s) \rightarrow 0 \tag{8}$$

Mean imputation aids in maintaining dataset integrity by replacing missing values with the mean of observed entries. Standard scaler normalization ensures uniform feature scaling, promoting algorithm convergence and model stability. Additionally, the use of isolation forests effectively identifies and eliminates outliers, contributing to the overall robustness of the predictive models. By integrating these preprocessing techniques, our study aims to provide a solid foundation for accurate and resilient data analysis, ultimately improving the quality of predictive insights derived from the models.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2025.3604710

B. Automated Decision Maker (ADM)

As described in Fig. 1, the ADM consists of automated pipelines of ML and DL models that work parallelly, in ensemble fashion, to define an optimized pipeline for the dataset. The main idea behind this pipeline is automating the model selection & hyperparameter tuning which consistently decreases the amount of time & resources spent on the hyperparameter tuning when done manually. This ADM works with the parallel collaboration of autoML (TPOT) & autoDL (AutoKeras) to identify an optimized supreme model for classification purposes.

Tree-based Pipeline Optimization Tool (TPOT) [26] is open-source Genetic Programming (GP) based AutoML system, that is built to optimize feature preprocessors and ML models and maximize supervised classification tasks. It can be summarized using,

$\Phi = (\prod D_c.FP_i.FC_i.KBest.Model_i).test(\tau_X, \tau_y)$

where Π, D_c, FP_i, FC_i, $KBest$, $Model_i$, i represents a series sequence, data copies, feature preprocessing, feature combination, selection of Kbest features, models, and the number of available options in TPOT algorithm. The process of TPOT is optimized using the GP, as described in Algorithm 2. It describes a GP-based (Genetic Programming) optimization approach for TPOT to find the best classifier. Initially, 100 random tree-based pipelines are generated and evaluated for cross-validation accuracy. The top 20 pipelines are selected using the NSGA-II selection scheme, focusing on accuracy and pipeline simplicity. These pipelines undergo genetic operations: 5 copies are made as offsprings, with 5% subjected to one-point crossover and the rest undergoing mutation with a 1/3 probability. The best solution is updated iteratively, and the optimal pipeline is returned. This method balances classification accuracy and pipeline efficiency.

AutoKeras simplifies the intricate task of constructing and training deep neural networks, providing an accessible interface for novice users. Built on Keras and TensorFlow, it facilitates solving standard ML problems with just a few lines of code. AutoKeras-designed models are seamlessly exportable and deployable using the TensorFlow ecosystem tooling, making it ideal for practical applications. AutoKeras is based on the following problem statement: Say a neural structure search domain N, with $T_{X_{train}}$ & $T_{X_{val}}$ be the training and validation data, with cost function ψ, it is aimed to find the optimal neural network $\eta^* \in$N with the lowest cost. The definition is equivalent to finding η^* satisfying:

$$\eta^* = argmin_{\psi}(\eta(p^*), T_{X_{val}}) \tag{9}$$

$$p^* = argmin_{\eta} L(\eta(p), T_{X_{train}}) \tag{10}$$

where p^* is the learned parameter of η. This search domain N is covering all neural structures. AutoKeras uses a search algorithm that is defined in Algorithm 3.

The hyperparameters are grouped into sub-modules and they are assigned probabilities $\hat{P}$ defined as,

$$\hat{P} = (\frac{1}{hp_1 + 1}, \frac{1}{hp_2 + 1}, ..., \frac{1}{hp_k + 1}) \in R^K \tag{11}$$

Algorithm 2 GP-based TPOT optimization algorithm

Input: Training data T & testing data τ
Output: Best TPOT Classifier

```
Generate 100 random tree-based pipelines say R_t.
Evaluate their cross-validation accuracy on the dataset.
for i in R_t:
    t_20=top 20 pipelines according to NSGA-II selection scheme [27]
    Produce 5 copies(t_20) [called offsprings]
    Crossover 5% of offsprings (using one-point crossover)
    Mutate rest copies (chances of mutation are 1/3)
    Best_sol^n = max(best_pipeline)
return Best_sol^n
```

Note: The pipelines are selected for maximizing classification accuracy and minimizing number of operators in the pipeline.

Algorithm 3 AutoKeras search algorithm

```
for i in e:
    if i ≤ c
        evaluate(i^th pre-defined hyperparameter)
    else
        evaluate(mutate(find_best_hyperparameter()))
```

where: e is total number of evaluations, c is number of predefined configurations

where hp_i is the hyperparameter of i^{th} submodule, K is number of submodules. To normalize $\hat{P}$, it is passed through logit and softmax function thus giving P as,

$$P = Softmax(logit(\hat{P})) = Softmax(-ln\, hp) \tag{12}$$

$$logit(a) = \frac{a}{1-a} \tag{13}$$

and hp = $(hp_1, hp_2, ..., hp_K \in R^K)$

IV. Results & Evaluation

This section details experimental results on UCI-HAR and PAMAP2 datasets, covering data characteristics, performance metrics, and obtained results.

A. Experimental Setup

The proposed framework was developed in Python using AutoKeras, and TPOT libraries where both the datasets had a 70:30 train-test ratio and training set was stratified using 10 splits and 3 repeats. The implementation and evaluations of the proposed framework were conducted using Python 3.10.11 on Asus Vivobook, with 16 GB RAM, 512 SSD, and a 1650-Nvidia dedicated graphics card.

B. Dataset Description

The proposed methodology is tested upon the UCI-HAR and PAMAP2 datasets. The UCI-HAR dataset [28] is widely used in HAR research, capturing activity data from 30 volunteers equipped with waist smartphones containing accelerometers, magnetometers, and gyroscopes. It consists of 7352 training samples and 2947 test samples, providing a comprehensive resource for activity recognition studies with Table I representing the labels given to the activities.

TABLE I
LABELS FOR UCI-HAR DATASET

Activity	Label
Walking	1
Walking upstairs	2
Walking downstairs	3
Sitting	4
Standing	5
Laying	6

The PAMAP2 [29] dataset serves as a valuable resource for monitoring physical activities, involving 9 subjects equipped with three IMUs and heart rate monitors. This dataset is instrumental for researchers interested in strength estimation and activity recognition. The activities and labels opted for it are described in Table II.

TABLE II
LABELS FOR PAMAP2 DATASET

Activity	Label
Lying	1
Sitting	2
Standing	3
Walking	4
Running	5
Cycling	6
Nordic walking	7
Ascending stairs	12
Descending stairs	13
Vacuum cleaning	16
Ironing	17
Rope jumping	24

C. Performance Metrics

When evaluating HAR, essential metrics include accuracy, F-score, precision, and Mean Absolute Error (MAE). Accuracy assesses the overall correctness. Precision quantifies the ratio of accurately classified values to the total samples identified as positive whereas F-score computes the harmonic mean of precision and recall.

$$Accuracy = \frac{p_t + n_t}{p_t + n_t + p_f + n_f} \tag{14}$$

$$Precision = \frac{p_t}{p_t + p_f} \tag{15}$$

$$F - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{16}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y - \hat{y}| \tag{17}$$

Where p_t is true positive, n_t is true negative, p_f is false positive, and n_f is false negative respectively.

D. Results Analysis

This section analyses the results obtained on the UCI-HAR and PAMAP2 datasets using the ADM methodology.

Table III provides a comprehensive state-of-the-art ML & DL approaches comparison with the proposed methodology on the UCI-HAR dataset. The evaluated classifiers include Random Forest (RF), k-nearest Neighbors (KNN), Decision Tree (DT), Multi-Layer Perceptron (MLP), Stochastic Gradient Descent (SGD), and Perceptron.

TABLE III
STATE-OF-ART ML & DL APPROACHES COMPARISON WITH THE PROPOSED METHODOLOGY ON UCI-HAR DATASET

Classifier	Recall	Precision	F-score	MAE
RF	92.498	92.655	92.481	0.0849
KNN	89.07	89.353	88.983	0.131
DT	85.743	85.792	85.691	0.155
MLP	94.841	95.028	94.934	0.055
SGD	94.942	95.260	94.925	0.056
Perceptron	94.671	95.025	94.663	0.058
Proposed	**96.436**	**96.502**	**96.427**	**0.0418**

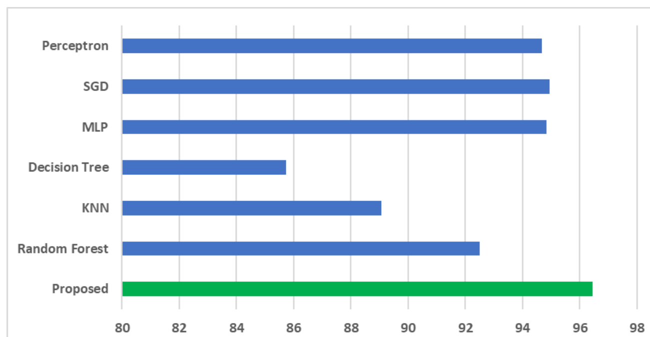

Fig. 2. Accuracy comparison on UCI-HAR dataset

TABLE IV
STATE-OF-ART ML & DL APPROACHES COMPARISON WITH THE PROPOSED METHODOLOGY ON PAMAP2 DATASET

Classifier	Recall	Precision	F-score	MAE
RF	99.640	99.640	99.639	0.028
KNN	99.51	99.511	99.509	0.012
DT	98.284	98.308	98.286	0.132
MLP	97.090	97.157	97.084	0.146
SVC	50.384	70.297	47.666	4.370
SGD	49.729	77.446	49.246	3.743
Perceptron	60.859	68.285	57.066	2.419
Proposed	**99.783**	**99.784**	**99.782**	**0.007**

The proposed methodology stands out with superior results across all metrics, giving a recall of 96.436%, precision of 96.502%, F-score of 96.427%, and MAE of 0.0418. In comparison to the SGD (second highest recall), ADM achieves approximately 2% increase value highlighting its ability to correctly identify the correct positives out of all positive instances.

Moerover, Fig. 2 displays accuracy values for various classifiers. The proposed methodology achieves an accuracy of 96.436%, outperforming other classifiers such as

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2025.3604710

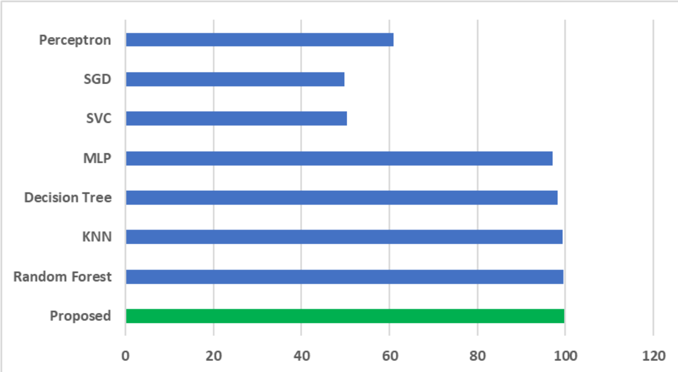

Fig. 3. Accuracy state-of-art comparison on PAMAP2 dataset

SGD (94.942%) which is the second highest, and Perceptron (94.671%) which is the third highest. These comparisons offer a detailed insight into the effectiveness of the proposed methodology in comparison to other well-known classifiers on the UCI-HAR dataset.

Another comparison of the proposed methodology on the PAMAP2 dataset against various classifiers is given in Table IV. The proposed methodology stands out with exceptional results, achieving recall of 99.783%, precision of 99.784%, F-score of 99.782%, and MAE of 0.007 with Fig. 3 representing accuracy values for different classifiers. In comparison to the RF with second highest metrics values, ADM achieves 75% better results in terms of MAE value. This signifies that the model prediction of the RG are slightly more off in comparison to the ADM methodology, highlightly its accuracy in detecting the activities near to the correct instances in comparison to the RF.

Fig. 4 summarizes cross-validation (CV) scores achieved using ADM methodology on the UCI-HAR dataset. Across generations, the CV score improved slightly, starting at 98.5263 for Generation 2 and increasing incrementally to 98.55 by Generation 10. This indicates a steady refinement of the model‘s performance through iterative optimisation.

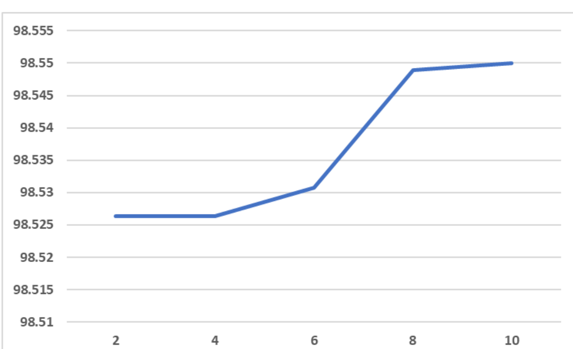

Fig. 4. Cross-validation score on UCI-HAR dataset with each generation

Likewise, Fig. 5 showcases cross-validation (CV) results on the PAMAP2 dataset. The CV score demonstrates consistent improvement, increasing from 99.099 at Generation 2 to 99.702 at Generation 8, where it stabilises through Generation 10. This trend resembles the pattern observed the UCI-HAR dataset, indicating the ability of the ADM methodology to remain unaffected with respect to the dataset. This could allow this approach to be extended and to be used other classification tasks such as anomaly detection or prediction.

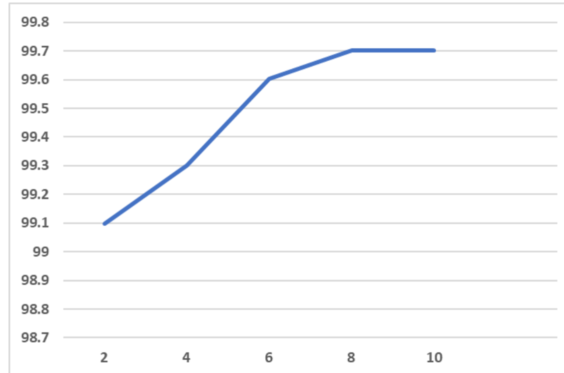

Fig. 5. Cross-validation score on PAMAP2 dataset with each generation

Fig. 6 describes the confusion metrics obtained through the ADM pipeline with Fig. 7 describing the classification report for the proposed methodology. It describes that the proposed methodology performs well in multi-class classification, getting higher support values for all classes. Likewise, Fig. 8 & Fig. 9 describe the classification report and confusion metrics of the proposed approach on the PAMAP2 dataset, showcasing the supremacy of the automated pipeline in correctly identifying all the classes. A careful observation of these classification reports indicates that, irrespective of the class (activity), ADM methodology successfully identifies the specific type of activity based on its determining features.

A comprehensive comparison of the ADM methodology with existing research works on the UCI-HAR dataset is provided in Table V. ADM consistently outperforms these works across multiple metrics, highlighting its effectiveness

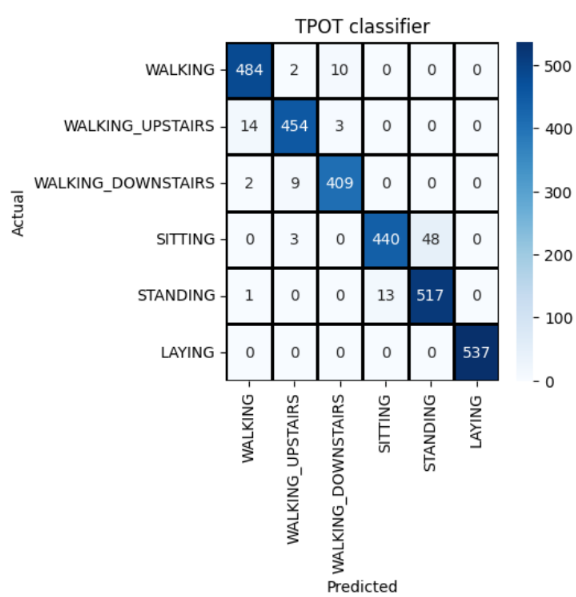

Fig. 6. Confusion metrics of the proposed approach for UCI-HAR dataset

```
              precision    recall  f1-score   support

           1       0.97      0.98      0.97       496
           2       0.97      0.96      0.97       471
           3       0.97      0.97      0.97       420
           4       0.97      0.90      0.93       491
           5       0.92      0.97      0.94       531
           6       1.00      1.00      1.00       537

    accuracy                           0.96      2946
   macro avg       0.97      0.96      0.96      2946
weighted avg       0.97      0.96      0.96      2946
```

Fig. 7. Classification report of the proposed approach on UCI-HAR dataset

```
              precision    recall  f1-score   support

           1       1.00      1.00      1.00      2945
           2       1.00      1.00      1.00      2734
           3       1.00      1.00      1.00      2916
           4       1.00      1.00      1.00      3539
           5       1.00      1.00      1.00      1447
           6       1.00      1.00      1.00      2470
           7       1.00      1.00      1.00      2852
          12       0.98      0.99      0.99      1774
          13       0.99      0.98      0.99      1576
          16       1.00      1.00      1.00      2561
          17       1.00      1.00      1.00      3622
          24       1.00      1.00      1.00       708

    accuracy                           1.00     29144
   macro avg       1.00      1.00      1.00     29144
weighted avg       1.00      1.00      1.00     29144
```

Fig. 8. The proposed methodology‘s classification report on PAMAP2 dataset

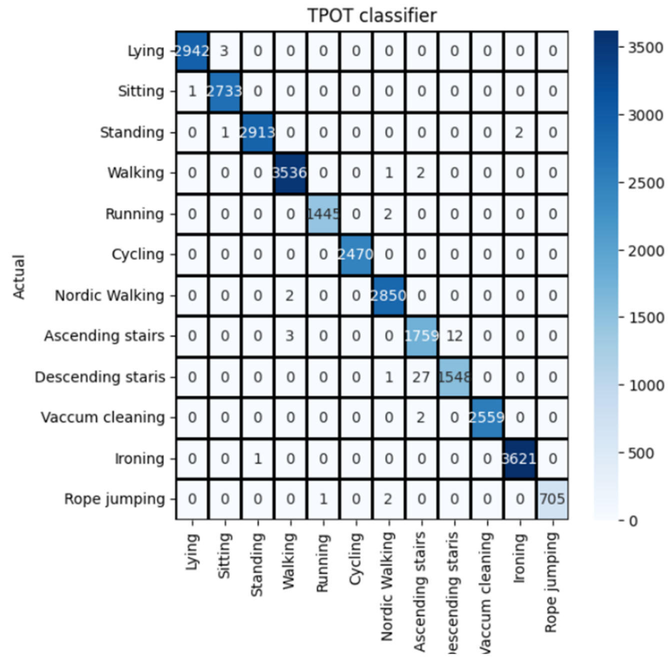

Fig. 9. The proposed methodology‘s confusion metrics on PAMAP2 dataset

in comparison to existing approaches.

TABLE V
COMPARING THE PROPOSED METHODOLOGY WITH EXISTING RESEARCH WORKS ON UCI-HAR DATASET

Approach	Accuracy	Precision	Recall	F-score
[23]	92.710	92.930	92.860	92.620
[30]	93.400	92.200	93.550	93.480
[31]	95.380	95.460	95.410	95.360
[32]	95.750			-
[33]	95.180	-	-	-
[34]	96.770	-	-	-
[35]	93.4	-	-	93.5
[36]	96.0	-	-	-
Proposed	**96.436**	**96.502**	**96.436**	**96.427**

Comparisons are made with notable works, including [23] with an accuracy of 92.71 and an F-score of 92.62, [30] with an accuracy of 93.4 and an F-score of 93.48, [31] with an accuracy of 95.38 and an F-score of 95.36, [32] with an accuracy of 95.75, [33] with an accuracy of 95.18, and [34] with an accuracy of 96.77.

Similarly, Table VI presents a comparison of the proposed methodology with existing research works. It demonstrates superior performance across all metrics, highlighting its effectiveness in comparison to existing approaches on the PAMAP2 dataset. As previously mentioned, ADM utilizes an automated pipeline, effectively eliminating the human errors commonly found in traditional and state-of-the-art approaches. The results obtained highlight and support this advantage.

TABLE VI
COMPARING THE PROPOSED METHODOLOGY WITH EXISTING RESEARCH WORKS ON PAMAP2 DATASET

Model	Accuracy	Precision	Recall	F-score
[37]	92.970	-	-	93.030
[32]	93.500	-	-	-
[38]	93.700	-	-	-
[39]	86.000	-	-	-
[40]	89.960	-	-	-
[34]	91.460	-	-	-
[41]	98.58	-	-	-
[42]	99.27	96.150	96.361	96.540
Proposed	**99.783**	**99.784**	**99.783**	**99.782**

V. DISCUSSION

In this section, we delve deeper into the implications of the experimental results obtained from the implementation of the ADM system in HAR. The experiments conducted and observed in the above section establish the efficacy of the proposed approach over existing state-of-the-art methodologies. The notable performance metrics observed on the UCI-HAR and PAMAP2 datasets, achieving accuracies of 96.436% and 99.783%, respectively highlight ADM's effectiveness in automating the HAR pipeline. These results signify a substantial improvement over traditional, manually designed approaches. This performance gain stems from ADM's ability to automatically perform parameter tuning and model selection, ensuring optimal classifier performance. In contrast, conventional methods depend heavily on manual configuration, which often limits their efficiency, scalability, and consistency.

This improvement is achieved because the ADM employs an automated process to tune parameters, ensuring the best-performing classifier. In contrast, other classifiers rely on manual tuning or user-provided parameters, which makes it challenging for them to achieve comparable performance levels due to the limitations of manual fine-tuning. These results highlight the transformative potential of ADM in enhancing the accuracy and efficiency of HAR systems, particularly in healthcare monitoring and sports analytics applications. Moreover, the results on two different datasets underscore the robustness of the proposed approach in controlled experimental conditions.

However, it is essential to address the real-world applicability of the ADM system. While the experimental results are promising, deploying ADM in real-time scenarios poses challenges that require further exploration. One such challenge is ensuring system performance with real-time data, where variability in sensor quality, environmental conditions, and human behavior can significantly affect outcomes. In addition, computational requirements for processing large-scale real-time data streams must be considered to ensure the scalability and efficiency of the system in practical applications.

Despite these challenges, ADM offers transformative potential by streamlining data processing and minimizing the need for human intervention. Its scalable architecture enables seamless integration into existing IoT ecosystems, paving the way for real-time HAR systems across diverse domains. In healthcare, ADM can facilitate continuous remote monitoring and timely alerts, enhancing patient safety and responsiveness. In sports, it enables detailed motion analysis to improve performance and prevent injuries. Industrial settings can benefit from ADM's ability to detect unsafe behaviors and preempt accidents, while smart home environments can leverage activity recognition for personalized automation and comfort. Furthermore, security systems can utilize ADM to enhance anomaly detection and threat identification.

Possible use cases for the proposed approach span several sectors. In healthcare, ADM supports remote monitoring and emergency alerts, providing timely interventions for patients. For sports, it refines training through detailed movement analysis, optimizing performance and reducing injury risks. Industrial applications emphasize safety by identifying hazards and preventing accidents. Smart homes benefit from tailored environmental controls for comfort and efficiency. Security systems leverage improved anomaly detection to bolster safety in various settings. These multifaceted applications highlight ADM's potential, but practical implementation requires addressing the identified challenges to ensure consistent performance and user trust.

VI. Conclusion & Future Work

This research introduces a pioneering methodology to HAR, representing a significant leap forward in the field. The proposed approach demonstrated the exceptional performance on widely used UCI-HAR and PAMAP2 datasets, and outperform state-of-the-art models. The proposed ADM model showed outstanding results on the UCI-HAR dataset, with an accuracy of 96.436%, precision of 96.502%, recall of 96.436%, and a notable F-score of 96.427%. On PAMAP2 dataset, our approach excelled with impressive accuracy, precision, recall, and F-score, and with a minimal MAE of 0.007. Notably, proposed approach brings about automation by eliminating the need for hyperparameter tuning, mitigating human errors, and streamlining the model development process. The removal of manual intervention in hyperparameter tuning enhances the reproducibility and reliability of the model, making it a robust and user-friendly solution. The novel contributions presented in this paper pave the way for continued exploration and innovation in the field, fostering a deeper understanding of human activity patterns and enhancing the practical applications of activity recognition technologies. However, despite the promising results and potential benefits and use cases of ADM, several challenges and limitations must be addressed in the future. Future work should focus on addressing these limitations to enhance the practical viability of ADM. Key areas include optimizing computational efficiency, developing adaptive algorithms to handle sensor variability, and ensuring robust performance in dynamic and less controlled environments. By tackling these challenges, ADM can fully realize its potential as an indispensable tool across healthcare, sports, industry, smart homes, and security.

Acknowledgment

This publication has emanated from research supported by a grant from Taighde Éireann - Research Ireland under grants 12/RC/2289_P2 (Insight) and 21/FFP-A/9174 (SustAIn). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

[1] K. Liu, C. Gao, B. Li, and W. Liu, "Human activity recognition through deep learning: Leveraging unique and common feature fusion in wearable multi-sensor systems," *Applied Soft Computing*, vol. 151, p. 111146, 2024.

[2] Z. Liu, J. Li, Z. Shen, G. Huang, S. Yan, and C. Zhang, "Learning efficient convolutional networks through network slimming," in *Proceedings of the IEEE international conference on computer vision*, pp. 2736–2744, 2017.

[3] J. Li, B. J. Cairns, J. Li, and T. Zhu, "Generating synthetic mixed-type longitudinal electronic health records for artificial intelligent applications," *NPJ Digital Medicine*, vol. 6, no. 1, p. 98, 2023.

[4] D. Hendry, K. Chai, A. Campbell, L. Hopper, P. O'Sullivan, and L. Straker, "Development of a human activity recognition system for ballet tasks," *Sports medicine-open*, vol. 6, no. 1, pp. 1–10, 2020.

[5] S. Tian, W. Yang, J. M. Le Grange, P. Wang, W. Huang, and Z. Ye, "Smart healthcare: making medical care more intelligent," *Global Health Journal*, vol. 3, no. 3, pp. 62–65, 2019.

[6] R. Zhang, C. Jiang, S. Wu, Q. Zhou, X. Jing, and J. Mu, "Wi-fi sensing for joint gesture recognition and human identification from few samples in human-computer interaction," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2193–2205, 2022.

[7] B. Anzengruber-Tanase, G. Sopidis, M. Haslgrübler, and A. Ferscha, "Determining best hardware, software and data structures for worker guidance during a complex assembly task," in *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 63–72, 2022.

[8] G. Sopidis, M. Haslgrübler, B. Azadi, B. Anzengruber-Tánase, A. Ahmad, A. Ferscha, and M. Baresch, "Micro-activity recognition in industrial assembly process with imu data and deep learning," in *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 103–112, 2022.

[9] M. Laube, M. Haslgrübler, B. Azadi, B. Anzengruber-Tánase, and A. Ferscha, "Skill level detection in arc welding towards an assistance system for workers," in *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 73–82, 2022.

[10] S. Wang, C. Wang, C. Zhang, J. Xue, P. Wang, X. Wang, W. Wang, X. Zhang, W. Li, G. Huang, *et al.*, "A classification-based spatiotemporal adaptive fusion model for the evaluation of remotely sensed evapotranspiration in heterogeneous irrigated agricultural area," *Remote Sensing of Environment*, vol. 273, p. 112962, 2022.

[11] C. Dobbins, R. Rawassizadeh, and E. Momeni, "Detecting physical activity within lifelogs towards preventing obesity and aiding ambient assisted living," *Neurocomputing*, vol. 230, pp. 110–132, 2017.

[12] J. Donahue, L. Anne Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell, "Long-term recurrent convolutional networks for visual recognition and description," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2625–2634, 2015.

[13] C. P. Burgos, L. Gärtner, M. A. G. Ballester, J. Noailly, F. Stöcker, M. Schönfelder, T. Adams, and S. Tassani, "In-ear accelerometer-based sensor for gait classification," *IEEE Sensors Journal*, vol. 20, no. 21, pp. 12895–12902, 2020.

[14] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *the Journal of machine Learning research*, vol. 9, pp. 1871–1874, 2008.

[15] 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, p. 322, 2022.

[16] H. Kang, Q. Huang, and Q. Zhang, "Augmented adversarial learning for human activity recognition with partial sensor sets," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 3, pp. 1–30, 2022.

[17] K. Muhammad, A. Ullah, A. S. Imran, M. Sajjad, M. S. Kiran, G. Sannino, V. H. C. de Albuquerque, *et al.*, "Human action recognition using attention based lstm network with dilated cnn features," *Future Generation Computer Systems*, vol. 125, pp. 820–830, 2021.

[18] D. Bhattacharya, D. Sharma, W. Kim, M. F. Ijaz, and P. K. Singh, "Ensem-har: An ensemble deep learning model for smartphone sensor-based human activity recognition for measurement of elderly health monitoring," *Biosensors*, vol. 12, no. 6, p. 393, 2022.

[19] C. Dai, X. Liu, and J. Lai, "Human action recognition using two-stream attention based lstm networks," *Applied soft computing*, vol. 86, p. 105820, 2020.

[20] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 115, pp. 213–237, 2019.

[21] Z. N. Khan and J. Ahmad, "Attention induced multi-head convolutional neural network for human activity recognition," *Applied soft computing*, vol. 110, p. 107671, 2021.

[22] F. Cruciani, A. Vafeiadis, C. Nugent, I. Cleland, P. McCullagh, K. Votis, D. Giakoumis, D. Tzovaras, L. Chen, and R. Hamzaoui, "Feature learning for human activity recognition using convolutional neural networks: A case study for inertial measurement unit and audio data," *CCF Transactions on Pervasive Computing and Interaction*, vol. 2, no. 1, pp. 18–32, 2020.

[23] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep learning models for real-time human activity recognition with smartphones," *Mobile Networks and Applications*, vol. 25, pp. 743–755, 2020.

[24] R. Mutegeki and D. S. Han, "A cnn-lstm approach to human activity recognition," in *2020 international conference on artificial intelligence in information and communication (ICAIIC)*, pp. 362–366, IEEE, 2020.

[25] S. Deep and X. Zheng, "Hybrid model featuring cnn and lstm architecture for human activity recognition on smartphone sensor data," in *2019 20th international conference on parallel and distributed computing, applications and technologies (PDCAT)*, pp. 259–264, IEEE, 2019.

[26] R. S. Olson and J. H. Moore, "Tpot: A tree-based pipeline optimization tool for automating machine learning," in *Workshop on automatic machine learning*, pp. 66–74, PMLR, 2016.

[27] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.

[28] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," in *Ambient Assisted Living and Home Care* (J. Bravo, R. Hervás, and M. Rodríguez, eds.), (Berlin, Heidelberg), pp. 216–223, Springer Berlin Heidelberg, 2012.

[29] A. Reiss and D. Stricker, "Introducing a new benchmarked dataset for activity monitoring," in *2012 16th international symposium on wearable computers*, pp. 108–109, IEEE, 2012.

[30] S. Deep and X. Zheng, "Hybrid model featuring cnn and lstm architecture for human activity recognition on smartphone sensor data," in *2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, pp. 259–264, 2019.

[31] Z. N. Khan and J. Ahmad, "Attention induced multi-head convolutional neural network for human activity recognition," *Applied Soft Computing*, vol. 110, p. 107671, 2021.

[32] Y. Tang, Q. Teng, L. Zhang, F. Min, and J. He, "Layer-wise training convolutional neural networks with smaller filters for human activity recognition using wearable sensors," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 581–592, 2020.

[33] W. Jiang and Z. Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks," in *Proceedings of the 23rd ACM international conference on Multimedia*, pp. 1307–1310, 2015.

[34] W. Huang, L. Zhang, Q. Teng, C. Song, and J. He, "The convolutional neural networks training with channel-selectivity for human activity recognition based on sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 10, pp. 3834–3843, 2021.

[35] A. Dubey and J. Zacharias, "Imu data based har using hybrid model of cnn & stacked lstm," in *2024 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI)*, pp. 1–6, IEEE, 2024.

[36] Y.-C. Lai, Y.-C. Kan, K.-C. Hsu, and H.-C. Lin, "Multiple inputs modeling of hybrid convolutional neural networks for human activity recognition," *Biomedical Signal Processing and Control*, vol. 92, p. 106034, 2024.

[37] Q. Teng, K. Wang, L. Zhang, and J. He, "The layer-wise training convolutional neural networks using local loss for sensor-based human activity recognition," *IEEE Sensors Journal*, vol. 20, no. 13, pp. 7265–7274, 2020.

[38] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," *arXiv preprint arXiv:1604.08880*, 2016.

[39] A. Khan, N. Hammerla, S. Mellor, and T. Plötz, "Optimising sampling rates for accelerometer-based human activity recognition," *Pattern Recognition Letters*, vol. 73, pp. 33–40, 2016.

[40] M. Zeng, H. Gao, T. Yu, O. J. Mengshoel, H. Langseth, I. Lane, and X. Liu, "Understanding and improving recurrent networks for human activity recognition by continuous attention," in *Proceedings of the 2018 ACM international symposium on wearable computers*, pp. 56–63, 2018.

[41] J. Yang, T. Liao, J. Zhao, Y. Yan, Y. Huang, Z. Zhao, J. Xiong, and C. Liu, "Domain adaptation for sensor-based human activity recognition with a graph convolutional network," *Mathematics*, vol. 12, no. 4, p. 556, 2024.

[42] A. Bollampally, J. Kavitha, P. Sumanya, D. Rajesh, A. Y. Jaffar, W. N. Eid, H. M. Albarakati, F. M. Aldosari, and A. A. Alharbi, "Optimizing edge computing for activity recognition: A bidirectional lstm approach on the pamap2 dataset," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18086–18093, 2024.