Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 10, No. 1, March 2022, pp. 177~186
ISSN: 2089-3272, DOI: 10.52549/ijeei.v10i1.3438

An Efficient Human Activity Recognition Model Based on Deep Learning Approaches

Aymen Jalil Abdulelah[1], Mohannad Al-Kubaisi[2], Ahmed Muhi Shentaf[3]
[1,3]Computer Engineering Techniques Department, Al-Maarif University College, AL Anbar, Iraq
[2]Computer Science Department, Al-Maarif University College, AL Anbar, Iraq

Article Info

Article history:

Received Oct 9, 2021
Revised Jan 26, 2022
Accepted Mar 21, 2022

Keyword:

Human Activity Recognition
Convolutional Neural Network
Deep Learning
Long Short-Term Memory

ABSTRACT

Human Activity Recognition (HAR) has gained traction in recent years in diverse areas such as observation, entertainment, teaching and healthcare, using wearable and smartphone sensors. Such environments and systems necessitate and subsume activity recognition, aimed at recognizing the actions, characteristics, and goals of one or more individuals from a temporal series of observations streamed from one or more sensors. Different developed models for HAR have been explained in literature. Deep learning systems and algorithms were shown to perform highly in HAR in recent years, but these algorithms need lots of computerization to be deployed efficiently in applications. This paper presents a HAR lightweight, low computing capacity, deep learning model, which is ideal for use in real-time applications. The generic HAR framework for smartphone sensor data is proposed, based on Long Short-Term Memory (LSTM) networks for time-series domains and standard Convolutional Neural Network (CNN) used for classification. The findings demonstrate that many of the deployed deep learning and machine learning techniques are surpassed by the proposed model.

Corresponding Author:

Aymen Jalil Abdulelah
Deputy of Head of Department
Computer Engineering Techniques Department
General Street, AL Ramadi City, AL Anbar, Iraq
Email: ayman.ja90@uoa.edu.iq

1. INTRODUCTION

HAR has been a hot subject over the past decade, as it is relevant in many fields such as healthcare, digital games, activities and general tracking systems. In addition, the ageing population is becoming one of the major worries of the nation. The population of over 65 years old was estimated to rise from 460 million by 2050 to 2 billion. This significant rise would have huge implications for social and health care. In the household, HAR is a powerful instrument in order to track the bodily, functional and cognitive health of elderly [1],[2].

The purpose of HAR is for human activities to be recognized in regulated and unregulated environments. Despite myriad implementations, HAR algorithms also face a lot of difficulties, including: sophistication and a spectrum of everyday tasks, intra-subject and inter-subject heterogeneity in a particular field of operation, trade-off between success and privacy, embedded and mobile system computing performance and the data annotation complexity [3]. Data is usually obtained through two major sources, environmental sensors and wearable devices, for training and evaluating HAR algorithms. In certain places, the environmental sensors may be ambient sensors, such as temperature sensors or video sensors. Wearable sensors are embedded in or integrated into devices or clothing, for example, smartphones and smartwatches [4].

HAR science has seen an explosion in deep learning (DL) approaches with respect to algorithmic execution, resulting in improved of the precision of identification and recognition. Deep learning (DL) approaches deliver very accurate outcomes in high-accuracy datasets, although the small size of the dataset, the lower dimension of the input data and expert information available in formulating problems can be used to further integrate Machine Learning (ML) models in many HAR applications. The growing interest in HAR can be correlated in all areas of everyday living, particularly in health and well-being applications, with the use of sensors and wearable appliances. This growing interest in HAR can be seen in the number of interested publications published in the past five years [5] .

Daily behaviors become lifelogging data if consistently collected for longer periods. When this behavior is recorded, there are long forecasts that need to be processed nearby. One of the most interesting technologies today is for storing the sensor data with the cloud to be processed and the amount of web services is open. However, the constant transfer of accelerometer signals to the cloud will lead to tremendous network traffic and increased latency. The detection of everyday operations on edge devices before they are sent to the cloud will remove this issue by reducing latency, expense, network traffic and time for response [6] [7].

This study provides an enhanced HAR edge system with a deep learning algorithm. In this article, there are a number of contributions. The first one defines the architecture of the proposed enhanced system. This system is constructed using Recurrent Neural Network (RNN) in combination with the deep learning algorithm Long Short-Term Memory (LSTM). Furthermore, the system is learned and evaluated using optimized parameters for six HAR operations on a resource-constricted edge system such as the Raspberry Pi4. The third is an experiment, which tests the efficiency of the proposed system, with 29 participants conducting six everyday activities: standing up, walking, jogging, seating, going down the stairs and going up the stairs. The fourth one, in the measurements of precision, accuracy, recall, f-measurement and confusion matrix, the results of the proposed system are compared with some models that have already been produced.

Accordingly, this paper is organized as follows. In section 2, the HAR background and literature review are produced. Section 3 presents proposed enhanced RNN-LSTM model. In section 4, system discerption and data collection are presented. The suggested model is compared to current models in Section 5. Section 6 contains a conclusion and recommendations for further work.

2. LITERATURE REVIEW

The HAR research's is using significant guideline in deep learning and IoT science. Research is being carried out in this direction in various practical applications such as health care, identification of gymnastically behavior, human collapse detection, etc. This section addresses numerous articles which are suggested systems for solving problems related to this subject.

Ronao *et al* suggested a profound neural network (convnets) for data-adaptive detection of human activity recognition (HAR) on a smartphone, using the accelerometer and the gyroscope. Contents not only take advantage of the inherent temporal local dependence of 1D time-series signals and the invariance of translation and hierarchical features of operations, they also provide a way to extract appropriate, robust features automated and data-adaptive without needing specialized preprocessing or time-consuming handcrafting features. Experiments illustrate that more complex functions are extracted in each other layer, but when the information approaches the top coevolutionary layers, the disparity in degree of complexity between neighboring layers is decreased. Another helpful thing is that a larger filter size is often shown, since the time local association between the adjacent sensors lasts longer. Furthermore, a reduced size is best implemented as the information transmitted from the input to the convolution/pool layers is very important to keep it intact [8]. Highlighting, classification and identification of human activities through smartphone sensors were the key problems of this paper. Time-series signals are depleted successfully during a convolution process and the combining process cancels the effects of subtle translations in the input. The characteristics are automatically derived from raw time series sensor data using a multilayer converter with different convolution and pooling layers, with less layer removing more simple features and larger layers resulting in more complicated ones.

Hassan *et al* have been recommended here to use inertial smartphone sensors such as actuators and gyroscope sensors to identify the operation. This framework primarily consists of three key components: detecting, retrieval of features, and identification. The portion is sensed. As inputs to the HAR system, it collects sensor data. Two popular sensors were chosen for data gathering on phones for this study: triaxial linear actuators and gyroscopes. The sensors have frequency statistics between 0 Hz and 15 Hz. The second significant component is the retrieval of features. This segment begins with noise reduction to separate relevant signals from triaxial acceleration, such as the acceleration. After eliminating the noise, statistics on fixed-size sliding windows are evaluated for the generation of robust features by using Kernel PCA to minimize dimension (KPCA). The third significant component of the framework is the modeling of functionality through deep learning by applying Deep Belief Network algorithm (DBN) and a Restricted Boltzmann Machines

(RBM) neural network as a pre-training process [9]. The main problems of this paper are the identification and measurement of human activities using smartphone sensors. There are behaviors such as cycling, biking, entertaining, exercise, etc. Based on the length and difficulty of the work: brief tasks, basic activities and complicated activities can then be separated into these three main categories.

The objective of the paper proposed by is to identify the most typical everyday activities of an individual home by incorporating a revolutionary HAR framework that uses the ability of the smart technology combined with the capabilities of deep learning. The developed wearable got connected an IMU and a Wi-Fi portion for transmitting data over a Cloud service provider and providing direct Internet access with a single home network to allows users to perform the installation process directly. The sensor is combined with the CNN network, which is configured to evaluate the limited resources available to allow cost effectiveness or wearable systems to maintain the direction it can be applied. The specific problem of this paper is the implementation of an overarching system to constantly track the actions of people in the field of environmental assistance, welfare administration, clinical issue, elderly treatment, recovery, entertaining and monitoring of smart home conditions. The cause of this interesting of this problem and it's testing in many applications, due to their scale and costs have steadily been reduced [10].

Another study by Wan *et al*, aims to present an infrastructure based on a phone inertial sensor for HAR. The proposed convolutional neural network (CNN) algorithm surpasses conventional solutions and displays state-of-the-art outcomes based on comparable studies. At the same point, HAR and deep learning approaches collaborate to increase calculating accuracy and quality to minimize energy consumption expense and equipment resources. It is a job of conduct identification for RGB-D or mobile data sources to communicate human body movements, to analyze the living stimuli and to recognize the semantics of behavioral events. After that, in human-like video data, the algorithm will comprehend and judge the action sequences Computers' precision and the forecast of behavioral behavior can be significantly promoted by overcoming the computational dimensionality and inconsistency of human conduct in health monitoring devices, smart houses, communications between human and machine, etc. The major point in this article is the design of HAR technologies for healthcare surveillance systems to resolve the differences among healthcare practitioners and patients. Integrated transportable sensors allow smartphones an all-round tool to capture and interpret data, allowing HAR to be done by smartphones. The challenge of identifying human behavior with the build-in mobile accelerometer has been solved well, but these conventional approaches do not distinguish complex and real-time human behaviors with multimodal and large data sources [11].

Gumaei *et al*, have suggested a system for the identification of human behavior based on a hybrid profound learning paradigm and multimodal body sensor data. The dual deep learning model involves a number of neural network layers that incorporate the two distinct types of recurring units (SRUs) and the Gated Recurrent Units (GRUs), which are referred to as deep SRU-GRU models. They used the advantage of quick, fast SRUs and more powerful and reliable GRUs. They also suggested the use of a dual deep learning model to merge simple and gated reciprocal neural network systems to provide a multi-sensor platform for human behavior detection. In addition, they utilize deep-simple recurrent devices to handle then use any internal memory capacities to handle the sequences of multi-sensors input information. This framework was used to refine the hyper parameters of the model on the basis of manual and grid search procedures. The proposed framework was tested using an MHEALTH multi-sensor dataset and the cutting-edge work on the same dataset was taken into account. The most popular deep learning techniques, however, fail to take into account pattern sequences or recall modifications to the pattern sequences over the duration of periods. Recurrent neural networks (RNNs), according to internal storage capacities, are used for solving these problems in many systems for impressive outcomes [12].

Li *et al*, have been suggested three hybrid deep learning algorithms to be used in there research in order to deal with Human Activity Recognition (HAR) problem. Power Spectral Density (PSD) parameters are directly derived prior to training the Deep Recurrent Neural Network (DRNN) model from vertical accelerations and triaxle accelerations that deliver an electrical power at a period along the time spectrum and thus thoroughly describe the features of human behavior across successive time stages. Furthermore, real number PSDs are chosen to increase measurement performance rather than sophisticated and complex numbers STFT output. Comparison of representative deep learning techniques for HAR problem using only the accelerometer sensor was performed to evaluate the suggested schemes in this article in detailed experimental tests based on the actual data set. A particular study has been made on the identification of complicated human activities related to the transition between the activities, and the final outcomes of experiments show that the proposed methods have good results in terms of result accuracy and implementation time. The study attempts to detect traditional human behaviors in the area of intelligent healthcare and has had fantastic applications. In HAR study, the identification of everyday life behaviors and activities and the complex dropping activities. Data collected by an accelerometer, the most common component on smartphones and wearables, are especially useful for offering a wealth of knowledge on human activity and can be used for handy HAR implementation. For these low-capacity instruments, productivity, efficiency and quality must be treated together for data classification [13].

Rav`ı *et al*, suggested a deep learning algorithm for the precise and significant identification of power-savvy smart technology and smartphones is developed for the identification of human behavior and attitude. A feature generation procedure implemented to the spectral domain of inertial data has been designed to prevent changes in sensor orientation, sensor location and sensor acquisition speeds. In particular, the approach proposed uses quantities of transformed input spatial convolutions. The method to a systemic HAR feature learning method was used in this article for taking advantage of the concept of deep learning. In order to make the convolutional layers invariant with a change in different characteristics, the proposed method takes a technique to retrieve information using a temporal convolution in the domain of the inertial data. The deep neural network includes links and relationships among input data, which are generally ignored. In addition, a wide range of layers for optimizing attribute design are placed on top of each other by attempting to capture all given the input combinations using a larger number of sensor nodes. Alterations in the sensor direction, sensor location on the human body, and other sensor configuration modifications make it difficult to derive productivity and helps from actual data. Where the process isn't appropriate for deeper learning, existing approaches for deep learning fix this by using a wide variety of frames and nodes, which results in higher model complexity, which isn't ideal for Body Sensor Networks (BSN) [14].

Tanberk *et al*, suggested a hybrid deep framework in their paper for the comprehension and analysis of videos based on recognizing human activities. The suggested platform is built using deep-learning techniques to combine dense flow technique and auxiliary knowledge in the video databases. In order to identify human behavior, 3D convolutional neuronal networks (3D-CNN) supplied to the optical flow were combined with long short-term memory (LSTM) supplied to contributory information through video stream. Furthermore, during the identification of clips, the assistance vector machine algorithm is used. The major point is how to build a complex interaction HAR hybrid-based video classification framework. In addition, in the consolidation phase how the hybrid deep model will classify videos and produce video subtitles [15].

Zebin *et al*, introduced a method of feature learning that uses deep convolutional neural networks (CNN) to systematically automate feature learning from the input variables. The effect on CNN output was monitored from various important hyperparameters, including several coevolutionary layers and kernel sizes. The studies also showed that Deep Convolutional Neural Networks (CNNs) can manipulate signal streams of gyroscope and accelerometer to auto-learn the optimal functionality of the input for the task of identification of operation. Data were gathered from five separate sensor sites on the lower part of the body to further identify the activities. In order to avoid prediction error, the neural network due to the smaller size of training, the researchers used techniques like pooling, reduction, and soft-max classification. The key challenge in this research was the achievement of high precision in the recognition of human behavior and reducing system costs [16].

Zhu *et al*, were providing a human activity recognition system using mobile phone accelerometer, gyroscope and magnetometer based on the convolution neural network (CNNs) was the designed system in this article. This article proposed a new optimization technique for CNN to overcome the ambiguity among extremely related behaviors such as going up and down the stairs. Several algorithms were used to create three system steps, these are pre-processing of input data, feature extraction, and selection of the features. There were two variants used in this article called CNN network of seven classes (CNN-7) and network of two classes (CNN-2). Two versions of CNN implement the metric pool to carry out human activities during the research process. They accepted that the ensemble learning approach is very successful in differentiating the uncertainty between certain very related and often confounding activities such as walking and going upstairs. CNN-7 identifies seven activities and CNN-2 distinguishes two confusing human actions that produce extremely similar signal patterns: going upstairs and walking. If the CNN-7 output does not go up or down, the output is the final decision. But if CNN-7 performance is up or down then combining the CNN-2 forecast to increase the precision of recognition of these two confounding activities was performed. The key challenge in this research was how to increase the accuracy and effectivity of HAR system, especially in some complex cases. These complex cases may be assumed as the difference in movements and walk-steps among people, the limited size of data, and the high complexity of individual actions [17].

Human Activities are encoded as series of sensor readings in time T. Traditional machine learning strategies carry out class mission with out shooting temporal co-relations among input samples. CNN addresses this problem, however, is confined via way of means of convolution kernel. RNN mixed with LSTM are capable of cope with this problem and feature gained recognition for growing HAR systems. In this paper, an optimized RNN-LSTM version which may be deployed on edge gadgets is developed.

3. PROPOSED METHOD

In conjunction with LSTM the suggested model is built using RNN. It has a superficial structure with only multiple convolutional layers and fifty neurons, making it possible to launch on edge computers such as IoT boarding devices (Raspberry Pi, Audrino etc.). The following is a summary of each part of the model.

3.1. Recurrent Neural Networks (RNN)

Timing information can be captured from sequence data by RNNs. It is a row of data layers. There are input, hidden and output layers. There are several nodes in the hidden layer and maybe multiple layers in it. Each hidden and output node have a producing function for the current hidden case and output value generation.

CNN is a feedforward neural network comprising hidden layers and input and output layers. Convolutional layers coupled with pooling layers produce hidden layers. Among the CNN's blocks, the convolution layer is the most important. In the convolutional layer, a convolution filter is used to input data in order to create a feature map that combines information from the filter. Multiple filters are applied to the input data in order to create a stack of feature maps, which modify the convolutional layer's final output. In the areas of original data, local dependencies are generated via a convolution technique. Furthermore, in order to associate non-linearity with the CNN, an additional activation function such as a rectified linear unit is added to feature maps.

A pooling layer, which stores the most important information, then reduces the total of samples in each feature map. With the use of the pooling layer, the training time was lowered while the dimensionality and over-fitting processes were minimized. The most common form of pooling function is max-pooling, which finds the biggest value in a given neighborhood window. CNN architectures are made up of convolutional and pooling layers, followed by a number of fully linked layers..

3.2. Long Short-Term Memory (LSTM)

The main issue of collapsing and disappearing gradients is faced by RNN networks. This prevents the network's ability to track a wide spectrum of time dependency for long background windows among data and human activities. This restriction can be eliminated from LSTM-based RNNs and long processing windows modelling by replacing conventional RNN nodes with LSTM memory unit. LSTM unit includes multiple variables and gates to monitor each cell's behavior. The activation functions of the gates monitor each activation function. The input data are supplied to various gates: forget gate, input gate, output gate, activation vector to get into activation function [18].

Via gates that transmit or block information through the LSTM unit, LSTM gets the capacity to make judgments about what to store and when to allow reads, writes, and deletions. LSTM is used to extract auxiliary information from hand tracking and movement based on chessboard recognition in this study.

3.3. Proposed RNN-LSTM System

The proposed system is started from a tri-axel accelerometer reading stage. The second stage is the input data window for converting data into an input vector. After that, the enhanced RNN-LSTM system will work on the input data and integrate the output from different states and give the prediction decision. At the last stage, the SoftMax classifier is used to produce the final results of that particular window. The proposed system involves of two hidden layers, each layer has fifty neurons [19].

The suggested LSTM network is used to learn characteristics, identify, and recognize events automatically. The LSTM network is a form of Recurrent Neural Network (RNN) that can solve time series correlation issues in both short and long time by employing the hidden layer as a memory cell. RNNs are a subset of ANNs that form a loop in internal networks. RNNs feature recurrent hidden states in which the output at each time step is dependent on the output at the previous time step, hence hidden cells in RNNs get feedback from prior states to present states [20].

Three gates control memory cell performance: input gate i_t, output gate o_t, and forget gate f_t. The following are the updating equations:

$$i_t = \text{sigmoid}\,(U_i h_{t-1} + W_i x_t + b_i)$$
$$o_t = \text{sigmoid}\,(U_o h_{t-1} + W_o x_t + b_o)$$
$$f_t = \text{sigmoid}\left(U_f h_{t-1} + W_f x_t + b_f\right)$$
$$\tilde{c}_t = \tanh\,(U_c h_{t-1} + W_c x_t + b_c)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$
$$h_t = o_t \odot \tanh\,(c_t)$$

The sigmoid function is employed in neural networks for prediction and transformation, and the operator ⊙ refers to vector element-wise multiplication.

However, the proposed model includes a unique RNN-LSTM-based ensemble that successfully resolves the misunderstanding between very comparable behaviors such as moving upstairs and walking. Extensive comparison tests utilizing traditional approaches, including RF, are undertaken to evaluate the performance of the ensemble model. The model uses the sum of rules to integrate output from several states using the SoftMax classifier to provide a single final output for that window.

4. SYSTEM AND DATA COLLECTION

Generally, a considerable quantity of training data is necessary for a multi-class classification issue, especially when the feature vector has a high dimension. Rich features extracted from a significant amount of training data can also successfully prevent overfitting and strengthen the model. This paper's data comes from a variety of sports scenarios with a variety of players and device locations. The data was gathered in such a way that the amount of data obtained for each action was virtually identical. We'll go through a lot of specifics about data gathering and Dataset in this part.

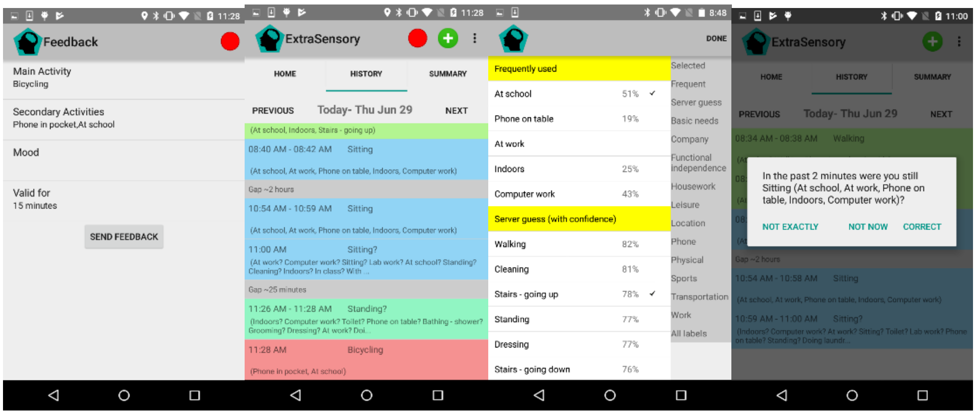
4.1. System Setup

The proposed system is designed on Raspberry Pi4 with 8 GB SDRAM and a processor of Quad-core Cortex-A72 (ARM v8), 64-bit, 1.5 GHz. The application of human activity recognition is applied on python v3.9 and TensorFlow deep learning library v2.4.0.

4.2. Dataset

The dataset for this work was collected online [21] and is called ExtraSensory. The dataset design proposed by Y. Vaizman and K. Ellis under the supervision of professor Gert Lanckriet. They published their paper [z], which describes the dataset in terms of user, device, sensor, and labels.

The author created an iPhone version and an Android version, as well as a Pebble watch component that works with both the iPhone and Android versions. For our dataset, we chose 30 samples. Furthermore, the Human Activity Recognition database was used to perform activities of daily living (ADL) while carrying a waist-mounted smartphone with embedded inertial sensors [22]. The built-in camera is used for collecting three-way data on Android devices. Six behaviors for 30 persons are included in the datasets. There are: standing up, walking, jogging, seating, going down the stairs and going up the stairs. Each behavior committed numerous acts in front of the leg pocket, holding a mobile phone. For the accelerometers, the persistent specimen rate of 20 Hz was calculated. Figure 1 depicts a screenshot of apps for collecting data sets as proposed by [21], followed by a comprehensive data set overview in Table 1.



(a) Active feedback (b) History view (c) Label selection (d) Notifications

Figure 1. Screenshot for colled dataset

Table 1. Summary of the dataset

Behavior	No. Samples	Ratio
Walking	4,24,400	38.6%
Jogging	3,42,177	31.2%
Going Upstairs	1,22,869	11.2%
Going Downstairs	1,00,427	9.1%
Seating	59,939	5.5%
Standing Up	48,397	4.4%

Enhanced RNN-LSTM is trained and separated into 60:40 for preparation and testing on a dataset. The models' weights are modified by the activation function. Middle cross-entropy between ground reality and forecast labels is used as a cost function. Using Adam Optimizer, costs can be minimized and model parameters modified. In Raspberry Pi4, this model has been learned to control the capacity of the system to interact with the tip. Various configurations have been tried and tested using the hit and trial approach to define limits, for example, a number of periods, the batch size, the window size, learning rate and kernel size. Table 2 shows the final set of selected hyperparameters.

Table 2. Model Hyperparameters

Hyperparameter	Value Selected
Number of Nodes	40
Number of epochs	100
Hidden Layers	4
Activation Function	Softmax
Optimizer	Adam
Learning Rate	0.0005
Time Step	100
Momentum	0.5

The commonly used test metrics are used to calculate the model's efficiency, precision, rendering, and f1 scores respectively.

I. Efficiency: Quantity of right forecasts for the complete number of TP sample forecasts.

$$Efficiency = (T + Tp)/(Tp + Tn + F + Fp) \quad (1)$$

where, Tp = true forecast of positive class, Tn = true forecast of a negative class, Fp = falsely positive forecast of class, Fn = falsely negative forecast of class.

II. Precision: Number of real forecasts about the system's overall true forecasts.

$$Precision = \frac{1}{c}\left(\sum_1^c \frac{Tp_c}{Tp_c+Fp_c}\right) \quad (2)$$

where, C = total amount of classes, Tp_c = true positives for a specific class c, Fp_c = falsely positive for a specific class c.

III. Rendering: Amount of real forecasts for the true number of system's predictions.

$$Rendering = \frac{1}{c}\left(\sum_1^c \frac{Tp_c}{Tp_c+Fn_c}\right) \quad (3)$$

where, C = total amount of classes, Tp_c = true positives for a specific class c, Fn_c = falsely negative for a specific class c.

IV. F1 scores: A harmonic average between precision and rendering calculation.

$$F1 = \sum_1^c 2\left(\frac{n_c}{N}\right) * \frac{Precision_c * Rendering_c}{Precision_c + Rendering_c} \quad (4)$$

where, N = total amount of samples, n_c = amount of samples in class c, Precision_c = precision amount for particular class c, $Rendering_c$ = rendering amount for particular class c.

5. EXPERIMENTAL RESULTS

This paragraph explains the enhanced RNN-LSTM evaluation findings and compares them to some current researches. 99% precision for jogging and walking was reached with the proposed enhanced RNN-LSTM. For going upstairs activities, a minimum precision of 81% is reached. Table 2 summarizes the outcomes of the evaluation metrics.

Table 2. System evaluation measurements

Algorithm	Efficiency	Precision	Rendering	F1 scores
RNN-LSTM	97.45 %	98.15 %	97.69 %	96.98 %

Calculation and assessment of performance, accuracy, recall, f1 ranking were applied. Enhanced RNN-LSLSTM reached 97.45% efficiency, 98.15% precision, 97.69% and 96.98% F1 scores. Efficiency will yield inaccurate results if the data in each data set class is unevenly retrieved and accurate output validation is measured.

The proposed method is compared with some of the others' methods: DL techniques, such as, CNN and its variations ([19], [15]), RNN ([18],[16]), ML techniques, such as, SVM ([17]), It perform better than them in terms of accuracy. The graphical representation of comparative results is shown in fig. 2.

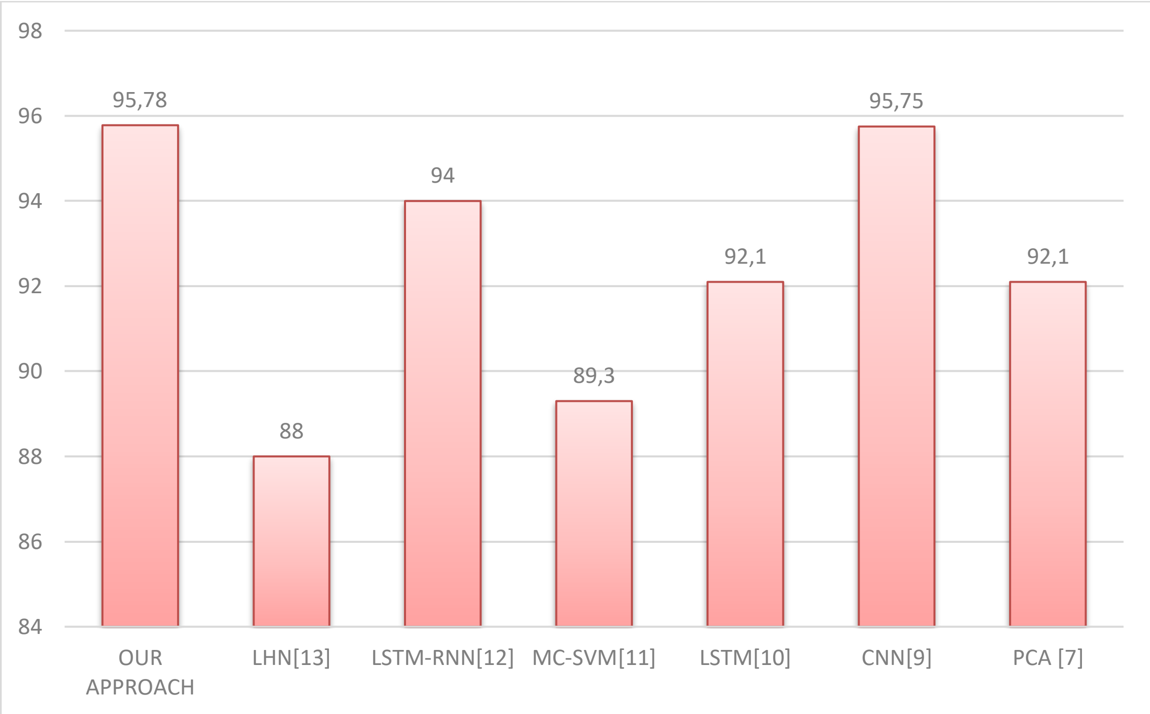


Figure 2. Comparison of accuracy with existing works

6. CONCLUSION

In the last decade, HAR systems have grown in popularity and made significant advances. Sensor-based HAR, in particular, has a number of benefits over vision-based HAR, which has privacy problems and is restricted by computing needs. Activity recognition techniques based on machine learning and deep learning are becoming more important in HAR. This paper develops an enhanced HAR model. This model is used on the Raspberry Pi 4 edge computer. The capture of human behaviors on edge devices decreases duration, expense and bandwidth utilization in networking. Particularly in comparison to several current machine learnings and deep learning systems, the developed system provides better outcomes.

With recent advancements in computer vision and deep learning, many experts believe that the difficulties in this field may be efficiently addressed and overcome. Different sensor types have their own set of benefits and drawbacks. During the design process, researchers should consider their options in light of the design goals. Combining complementing sensor categories can occasionally improve performance and give extra information, allowing them to transcend their respective limits

REFERENCES

[1] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, "Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities," *ACM Computing Surveys*. 2021, doi: 10.1145/3447744.

[2] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human Activity Recognition Using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey," *IEEE Access*, vol. 8, no. Dl, pp. 210816–210836, 2020, doi: 10.1109/ACCESS.2020.3037715.

[3] B. Fu, N. Damer, F. Kirchbuchner, and A. Kuijper, "Sensing technology for human activity recognition: A comprehensive survey," *IEEE Access*, vol. 8, pp. 83791–83820, 2020, doi: 10.1109/ACCESS.2020.2991891.

[4] Z. N. Khan and J. Ahmad, "Attention induced multi-head convolutional neural network for human activity recognition," *Appl. Soft Comput.*, 2021, doi: 10.1016/j.asoc.2021.107671.

[5] P. Agarwal and M. Alam, “A Lightweight Deep Learning Model for Human Activity Recognition on Edge Devices,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 2364–2373, 2020, doi: 10.1016/j.procs.2020.03.289.

[6] X. Zhou, W. Liang, K. I. K. Wang, H. Wang, L. T. Yang, and Q. Jin, “Deep-Learning-Enhanced Human Activity Recognition for Internet of Healthcare Things,” *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6429–6438, 2020, doi: 10.1109/JIOT.2020.2985082.

[7] E. Kim, S. Helal, and D. Cook, “Human activity recognition and pattern discovery,” *IEEE Pervasive Comput.*, 2010, doi: 10.1109/MPRV.2010.7.

[8] C. A. Ronao and S. B. Cho, “Human activity recognition with smartphone sensors using deep learning neural networks,” *Expert Syst. Appl.*, vol. 59, pp. 235–244, 2016, doi: 10.1016/j.eswa.2016.04.032.

[9] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, “A robust human activity recognition system using smartphone sensors and deep learning,” *Futur. Gener. Comput. Syst.*, vol. 81, pp. 307–313, 2018, doi: 10.1016/j.future.2017.11.029.

[10] V. Bianchi, M. Bassoli, G. Lombardo, P. Fornacciari, M. Mordonini, and I. De Munari, “IoT Wearable Sensor and Deep Learning: An Integrated Approach for Personalized Human Activity Recognition in a Smart Home Environment,” *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8553–8562, 2019, doi: 10.1109/JIOT.2019.2920283.

[11] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, “Deep Learning Models for Real-time Human Activity Recognition with Smartphones,” *Mob. Networks Appl.*, vol. 25, no. 2, pp. 743–755, 2020, doi: 10.1007/s11036-019-01445-x.

[12] A. Gumaei, M. M. Hassan, A. Alelaiwi, and H. Alsalman, “A Hybrid Deep Learning Model for Human Activity Recognition Using Multimodal Body Sensing Data,” *IEEE Access*, vol. 7, pp. 99152–99160, 2019, doi: 10.1109/ACCESS.2019.2927134.

[13] X. Li, Y. Wang, B. Zhang, and J. Ma, “PSDRNN: An Efficient and Effective HAR Scheme Based on Feature Extraction and Deep Learning,” *IEEE Trans. Ind. Informatics*, vol. 16, no. 10, pp. 6703–6713, 2020, doi: 10.1109/TII.2020.2968920.

[14] D. Ravi, C. Wong, B. Lo, and G. Z. Yang, “A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices,” *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 1, pp. 56–64, 2017, doi: 10.1109/JBHI.2016.2633287.

[15] S. Tanberk, Z. H. Kilimci, D. B. Tukel, M. Uysal, and S. Akyokus, “A Hybrid Deep Model Using Deep Learning and Dense Optical Flow Approaches for Human Activity Recognition,” *IEEE Access*, vol. 8, pp. 19799–19809, 2020, doi: 10.1109/ACCESS.2020.2968529.

[16] T. Zebin, P. J. Scully, and K. B. Ozanyan, “Human activity recognition with inertial sensors using a deep learning approach,” *Proc. IEEE Sensors*, no. 1, 2017, doi: 10.1109/ICSENS.2016.7808590.

[17] R. Zhu *et al.*, “Efficient Human Activity Recognition Solving the Confusing Activities Via Deep Ensemble Learning,” *IEEE Access*, vol. 7, pp. 75490–75499, 2019, doi: 10.1109/ACCESS.2019.2922104.

[18] A. A. Abdelsalam, A. A. Salem, E. S. Oda, and A. A. Eldesouky, “Islanding Detection of Microgrid Incorporating Inverter Based DGs Using Long Short-Term Memory Network,” *IEEE Access*, vol. 8, pp. 106471–106486, 2020, doi: 10.1109/ACCESS.2020.3000872.

[19] Y. Yu, X. Si, C. Hu, and J. Zhang, “A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures,” *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019, doi: 10.1162/neco_a_01199.

[20] J. Lei, C. Liu, and D. Jiang, “Fault diagnosis of wind turbine based on Long Short-term memory networks,” *Renew. Energy*, vol. 133, pp. 422–432, Apr. 2019, doi: 10.1016/j.renene.2018.10.031.

[21] Y. Vaizman, K. Ellis, and G. Lanckriet, “Recognizing Detailed Human Context in the Wild from Smartphones and Smartwatches,” *IEEE Pervasive Comput.*, vol. 16, no. 4, pp. 62–74, Oct. 2017, doi: 10.1109/MPRV.2017.3971131.

[22] F. Cruciani *et al.*, “A public domain dataset for human activity recognition in free-living conditions,” *Proc. - 2019 IEEE SmartWorld, Ubiquitous Intell. Comput. Adv. Trust. Comput. Scalable Comput. Commun. Internet People Smart City Innov. SmartWorld/UIC/ATC/SCALCOM/IOP/SCI 2019*, no. April, pp. 166–171, 2019, doi: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00071.

[23] H. Li and M. Trocan, “Deep learning of smartphone sensor data for personal health assistance,” *Microelectronics J.*, vol. 88, pp. 164–172, 2019, doi: 10.1016/j.mejo.2018.01.015.

[24] Ankita, S. Rani, H. Babbar, S. Coleman, A. Singh, and H. M. Aljahdali, “An efficient and lightweight deep learning model for human activity recognition using smartphones,” *Sensors*, vol. 21, no. 11, 2021, doi: 10.3390/s21113845.

[25] L. Dhammi and P. Tewari, “Classification of human activities using data captured through a smartphone using deep learning techniques,” 2021, doi: 10.1109/ICSPC51351.2021.9451772.

[26] A. J. Abdulelah, K. Shaker, A. M. Sagheer, and H. A. Jalab, “A dynamic scatter search algorithm for solving traveling salesman problem,” *Lect. Notes Electr. Eng.*, vol. 398, pp. 117–124, 2017, doi: 10.1007/978-981-10-1721-6_13.

[27] Y. A. Hamad, M. E. Seno, M. Al-Kubaisi, and A. N. Safonova, “Segmentation and Measurement of Lung Pathological Changes for COVID-19 Diagnosis Based on Computed Tomography,” *Period. Eng. Nat. Sci.*, 2021, doi: 10.21533/pen.v9i3.2086.

[28] N. S. Khan and M. S. Ghani, “A Survey of Deep Learning Based Models for Human Activity Recognition,” *Wireless Personal Communications*. 2021, doi: 10.1007/s11277-021-08525-w.

[29] E. Ramanujam, T. Perumal, and S. Padmavathi, “Human Activity Recognition with Smartphone and Wearable Sensors Using Deep Learning Techniques: A Review,” *IEEE Sensors Journal*. 2021, doi: 10.1109/JSEN.2021.3069927.

[30] A. S. Mustafa, A. J. Abdulelah, and A. K. Ahmed, “Multimodal biometric system iris and fingerprint recognition

based on fusion technique," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 3, pp. 7423–7432, 2020.

[31] A. L. Haroon P S and P. D. R. Premachand D. R, "Human Activity Recognition using Machine Learning Approach," *J. Robot. Control*, 2021, doi: 10.18196/jrc.25113.

[32] L. Alawneh, T. Alsarhan, M. Al-Zinati, M. Al-Ayyoub, Y. Jararweh, and H. Lu, "Enhancing human activity recognition using deep learning and time series augmented data," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 12, pp. 10565–10580, 2021, doi: 10.1007/s12652-020-02865-4.

[33] X. Li, P. Zhao, M. Wu, Z. Chen, and L. Zhang, "Deep learning for human activity recognition," *Neurocomputing*. 2021, doi: 10.1016/j.neucom.2020.11.020.

BIOGRAPHY OF AUTHORS

Aymen Jalil ABDULELAH Received his Master of Computer Science from the University of Anbar, Iraq. Currently serving as a lecturer in the Department of Computer Engineering Techniques at Al-Maarif University College. Also, he is a PhD student at Karabuk University, Karabuk, Turkey. He is highly interested in Computer Science, Image Processing, Network security and Artificial Intelligence. He can be contacted at email: ayman.ja90@uoa.edu.iq.

Mohannad Al-Kubaisi is currently a lecturer at the Computer Science department, AL-Maarif University college, Iraq. He holds an M.Sc. degree from The University of Anbar, Iraq in 2017 and received his Bachelor Degree in Computer Science from the University of Anbar, Iraq in 2013. He is highly interested in Digital Image Processing (DIP), Biomedical Signal Processing, artificial intelligence, deep learning and machine learning. He has published many papers in local and international journals and conferences. He can be contacted at email: mohannad@uoa.edu.iq.

Ahmed Muhi Shentaf is currently a lecturer at the Computer Engineering Technique department, AL-Maarif University college, Iraq. He holds an M.Sc. degree from The Altinbas University, Turkey in 2020 and received his Bachelor Degree in Computer Engireeing from AL-Maarif University College, Iraq in 2012. He is highly interested in Digital Image Processing (DIP), Ad-hoc Network, security and the design Algorithms in Network. He has published many papers in local and international journals and conferences. He can be contacted at email: ahmed.shantaf@uoa.edu.iq