JOURNAL OF NORTHEASTERN UNIVERSITY

Volume 26 Issue 01, 2023 ISSN: 1005-3026 https://dbdxxb.cn/ Original Research Paper

OPTIMIZING PHYSICAL ACTIVITY RECOGNITION USING HYBRID LSTM NETWORK

A. Ramya[1], C. Shoba Bindu[2], P. Dileep Kumar Reddy[3]

[1]Department of CSE, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India ramya999777@gmail.com

[2]Department of CSE, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India, shobabindhu.cse@jntua.ac.in

[3]Department of CSE, Narsimha Reddy Engineering College (Autonomous), Secunderabad, Telangana, India, dileepreddy503@gmail.com

ABSTRACT

Human Activity Recognition plays a crucial role in society. Due to the quick advancement of sensors such as smartwatches and other wearable devices recognizing human actions (HAR) have recently become more popular. The findings of significant HAR research projects are currently being applied in several mobile apps, such as health monitoring and athletic performance tracking, among others. Mainly, Human movement detection using sensors allows to predicting a person's movements using sensor-generated time- series data. In this paper, a HAR framework that uses automatically generated and is proposed to extract spatial-temporal information from data from smartwatch sensors. and also, we propose a Long Short-Term Memory Network (LSTM) and the 2D Convolutional Neural Network (2D-CNN) are employed in the framework to implement the hybrid deep learning approach, doing away with the necessity for feature extraction by manually. In this approach recognize the activities by considering both the significance of incoming short-term sensor data and the continuation of previous long-term sensor data activities. The outcomes showed the proposed deep learning hybrid LSTM model is more effective than the baseline models with an accuracy level of 91%. With an average improvement of more than 6% on the accuracy over the prior most effective model, we show that our suggested architecture that focused on attention is significantly more effective than prior approaches.

INDEX TERMS: Wearable devices, Accelerometer, Long Short-Term Memory, 2D Convolutional Neural Network, Deep Learning, Regular Daily Activities and Activity Recognition

I. INTRODUCTION

A time series classification challenge called SENSOR-BASED monitoring of human activities (HAR) involves tracking a person's movements or action in advance (such as walking, standing, running, etc.) based on sensor data. Gesture recognition, video surveillance, and fitness tracking are practical applications for HAR. HAR has been a particularly active study

Submitted: 11/12/2022 **Accepted**: 06/01/2023

field recently despite being a well- studied and mature subject because of the growth of ubiquitous computing made feasible by Internet-of-Things, wearable, and smartphone devices [1],[21]. A person's bodily movement can be detected at any time and everywhere using sensors like accelerometers, barometers, global positioning systems (GPS), gyroscopes. The human activity recognition problem is quite "personal," in that only one person typically uses a single smartphone or smartwatch, and each person has their own distinct motion when running, walking, or ascending stairs as discussed in fig 1. Deep learning methods that may be customized for a particular user are then preferred.

Numerous helpful mobile applications have embraced the benefits of wearable sensors, including detecting abnormal driving, remotely accessible health care systems monitoring elder people, tracking athletic performance, and mobile assistance systems for individuals who have vision problems.

Due to advancements in health [2], there is currently a bigger percentage of old people in the world's population than there ever was previously. Thus, there is a greater need for the care of the physical and mental health of individuals who live alone at the societal level. There's reason to be optimistic about AI and machine learning's ability to recognize tasks.

For seniors who want to age in place, activity recognition (AR) might be used to monitor their well-being, identify any concerning changes in routine [3], and notify helpers immediately if an emergency arises. Augmented reality (AR) may be subdivided into three categories according to the hardware used to collect data: camera video, wearable devices, and binary sensors. Camera video and wearable gadgets are less than perfect options because of privacy invasion worries and practical challenges, such as discomfort from the device and increased maintenance needs. To probe data-driven AR based on deep learning, this research devised a device-free, privacy-protecting method. A solution to the challenge of long-term activity monitoring in the real world was found in the binary sensor-based technique.

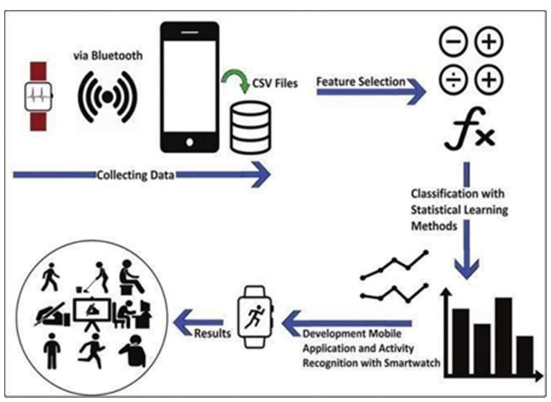


Figure 1. Human activity recognition (HAR) [3]

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

The AR process is not complete without the representation and extraction of features. This study set out to extract a meta-action by assessing the causal effect between a set of sensor activations in order to appropriately categorize and distinguish activities that are often conflated, such as standing, sitting, laying and walking etc. That's because every person's actions reflect their own set of beliefs, norms, and habits making human activity a process variable.

The unique sequence or feature of sensor activation in any given activity may be impacted by a person's habits and lifestyle, and this variance may be defined as a causality between sensors, even if the activity regions are similar. If these specifics can be identified and understood, it may be possible to enhance AR's performance.

Additionally, machine learning techniques may be used for advancements that can enhance the activity detection model employing smart devices to deliver more accurate evaluations of a wide range of activities. However, these methods employing standard machine learning frequently rely on automated feature extraction using heuristics, and are therefore typically restricted to the knowledge of the human domain. Due to this restriction, there are limitations on how well systems utilizing traditional machine learning perform in terms of the classification accuracy and other assessment metrics. In order to overcome this restriction, methods employing deep learning (DL) [4] are used in this paper.

This study proposed a HAR framework that applies a hybrid deep learning model is referred as a 2D CNN-LSTM network [6] to extract spatial-temporal [5] characteristics from data acquired by smart watch sensors. In order to compare the proposed hybrid LSTM-based strategy with the standard model of deep learning identified in a HAR dataset, an experimental analysis is carried out in this study. Wearable sensors are becoming a typical tool for business and professional applications. In actuality, contemporary smartphones and smartwatches come with sensors that enable the tracking and prediction of physical activity as well as the monitoring of physiological data.

The fall detection feature, which uses 3D time series data retrieved by an accelerometer [7] to determine whether a person has fallen and requires help, is a useful illustration of HAR. Sensors often gather data from a multidimensional time series in HAR, which poses significant issues.

- Noise: Sensor data is frequently noisy.
- Heterogeneous sensing rates: Various sensors might have various monitoring rates.
- Generalization and adaption of user: Everybody moves differently whether they run, jump, walk, etc.

In this study, 1) The HAR based on smartwatches focused on evaluating human behavior using sensors from smartwatches or smartphones, such as the gyroscope, accelerometer, etc. 2) Our goal is to create a model that accurately categorizes which of these activities is being done

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

using sensor data collected from research participants completing six distinct tasks. 3) This human activity recognition kept forward a wide range of purposes and advantages. The aged or senior assistance may benefit from this mobile-based health application. We also wish to utilize this application to monitor our personal health as it monitors our activities through time and is connected to our mobile device. 4)To increase the accuracy.

The goals are:

- To analyze the physical movement, Utilize the smartwatch sensors like the gyroscope, accelerometer, etc. Sensors use recognition to process the understanding while they go about their daily routines.
- Performing six different activities by using the sensor data through testing and comparing with different learning algorithms. we propose a hybrid deep learning model to achieve the best performance.
- Using objective support, we show effective method by demonstrating that it results in an average improvement of 6% on the accuracy on the predictions for a particular user. We also show that it works with every model and dataset and we look at.

The arrangement of the respective sections of this paper is as follows. The study on sensor-based activity identification employing deep learning and with some of the earlier papers that have been published presented in Section II. Information on the proposed LSTM-2D CNN architecture is described in Section III. Three public datasets for the implementation of network are provided in Section IV. Section V presents the experiment results and is also explained how model performance is influenced by network structure and hyper-parameters. This study's summary and conclusion are presented in the last section.

II. RELATED WORK

A. HUMAN ACTIVITY USING MACHINE LEARNING ALGORITHMS

Many approaches for modeling and identifying human activities have been presented in recent years as a result of the extensive study that has been done by researchers in investigating various sensing technologies. A. Jain, K. Gandhi, D. K. Ginoria and P. Karthikeyan [8] To categorize the data gathered by sensors, early studies mostly employed support vector machines (SVM), decision trees, and naive Bayes, and other conventional machine learning techniques. In A. Jain and V. Kanhangad [9], the characteristics of the angular velocity and acceleration data were extracted using Fourier descriptor and gradient histogram based on centroid feature. Then Jain et al. [8] employed the support vector machine and k-nearest neighbour (KNN) classifiers to identify the two activities open datasets. A total of six inertial measuring units were utilized to build a sensing device by Jalloul et al. [10]. The authors first conducted network analysis, after choosing a number of network measures that pass the statistical test to construct a feature set, the last stage was to use the random forest (RF) classifier to categorize the activities. The wearable wireless accelerometer-based activity identification system and its use

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

during medical detection were reported in the publication. For feature selection, Relief-F and sequential forward floating search (SFFS) were used. Finally, for activity identification and comparison analysis, k-nearest neighbour (KNN) and Naive Bayesian [11] K. Ashwini, R. Amutha and S. Aswin raj et al techniques were employed.

In the majority of everyday detecting the human action tasks, machine learning models could heavily rely on manual feature extraction using heuristics. Human domain knowledge typically places limitations on it. In order to address this issue, researchers have moved to deep learning techniques that, during the training phase, automatically extract the necessary features from unprocessed sensor data and show the abstract sequences of a high level over the original low-level temporal information. [12] Transferring deep learning models to the field of human activity detection is a new research path in pattern recognition in light of its successful use in the areas of speech recognition, natural language processing, and picture classification, and other domains. Three-axis accelerometer data was proposed to be converted into a "picture" format in [13], and human activities were then identified three convolutional layers and one fully-connected layer of CNN are used. (P. Agarwal and M. Alam et al) [14]. Although the models mentioned above might in general identify human activity, the total network structure is rather complex. These models also include a lot of parameters, which increases the cost of computing. When high real-time performance is required, it is challenging to employ. Numerous researchers have worked very hard in this area.

B. DEEP LEARNING ALGORTHMS FOR HAR

A lightweight deep learning model for HAR was developed by Agarwal et al. [15] and implemented on a Raspberry Pi3.This model was created by combining the LSTM algorithm and a shallow RNN. Even though the suggested model is quite accurate and has a simple architecture, although it was only one dataset with only six actions was assessed, which does not show how well the suggested model may be generalized. In the paper [16], a deep learning the combination of inception and the model (InnoHAR) for classifying activities, neural networks and recurrent neural networks are used. Separate convolution was employed by the authors to replace the conventional convolution, which was successful in its purpose of model settings. The findings show a wonderful impact; however, it took a long time for the model to barely converge throughout the learning phase.

Davide buffelli and Fabio vandin et al, [17] a novel deep learning framework, TrASenD, based on the state-of-the-art is overcome by our innovative deep learning framework, TrASenD, which is based only on a mechanism based on attention. With an average improvement of more than 7% in accuracy over the prior best-performing model, we show that our suggested architecture based on attention is far more effective than prior methods. We also take into account the issue of customizing HAR deep learning models, which is crucial in many applications.

To adapt a model to a particular user, we offer a straightforward and practical transfer learning-based approach that, on average, improves prediction accuracy for that user by 6%. With this

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

model, the average accuracy is 84%. The LSTM-2D CNN is a novel deep neural network for the recognition of human activity that can be used to overcome the drawbacks of the proposed techniques. The activity parameters could be automatically extracted by the model, and they could be quickly classified. Three of the most popular public datasets were also utilized to test it. This study showed that the proposed approach has excellent accuracy, good generalizability, and fast convergence speed.

III. PROPOSED ARCHITECTURE

As shown in Fig. 2, The collection of sensor data from the smartwatch sensor is achieved in the structure of human action recognition that is based on smartwatches in order to classify the activities carried out by smartwatch users. In this supervised machine learning techniques are used to create a labelled dataset. By using hybrid deep model, to train and test our labelled dataset and evaluating the training model. After achieving an acceptable level of validation accuracy and testing accuracy, we predict the physical movements of the person through the smart devices.

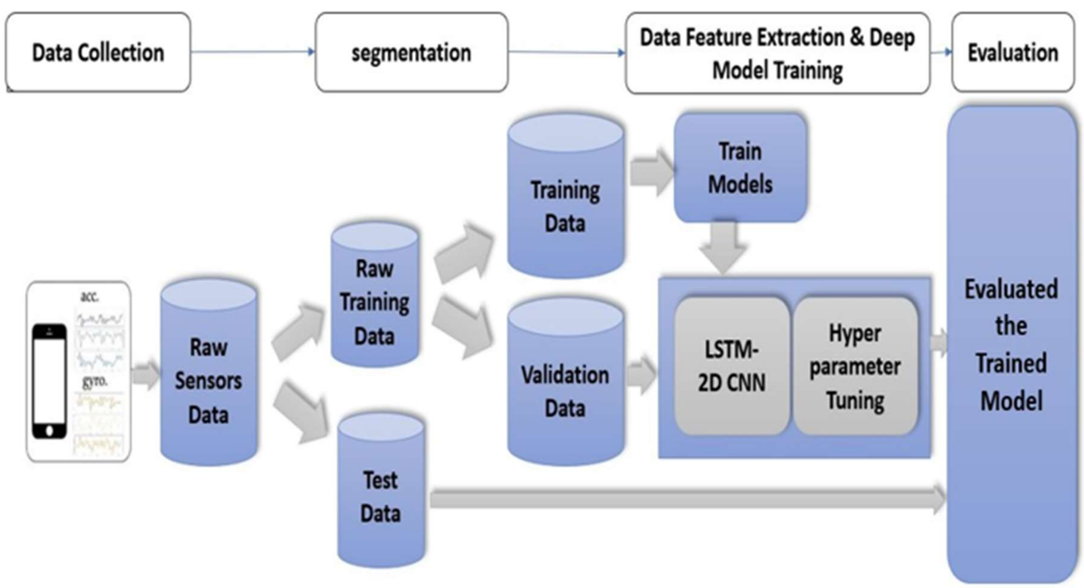


Figure 2. Workflow of hybrid LSTM model

A. SENSORS BASED: ACCELEROMETER

As shown in Fig 4 as depicted by Chandan Kashyap and Chandrashekhar et al, A smartphone's built-in accelerometer [18] is used to measure acceleration. Because it is three-dimensional, it can measure acceleration along the first, second, and third axes (As it is 3 dimensional). Both static and dynamic objects might be the subjects of the measurements. These days, all cell phones have these sensors, making their use both dependable and affordable.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

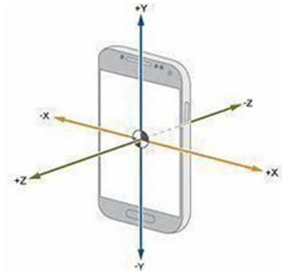


Figure 3. Accelerometer [18]

B. SENSORS BASED: GYROSCOPE

As shown in Fig 4 as depicted by Chandan Kashyap and Chandrashekhar et al, Gyroscopic sensors can be used to determine or maintain an object's angular velocity and orientation, whether it is static or moving [19]. This sensor, also known as a gyro-meter, can be found in cell phones in the form of a microchip- packed device. When readings of various human actions, such as walking, standing, sitting, etc., this sensor helps in maintaining stability.

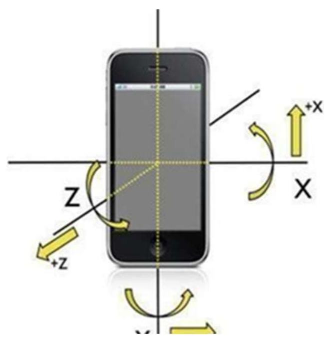


Figure 4. Gyroscope [18]

C. SENSOR BASED: PEDOMETER

M. Adjeisah, G. Liu, D. O. Nyabuga and R. N. Nortey [23], A pedometer is a tool that tracks a person's movement while they walk and counts their steps. Fitness enthusiasts now frequently utilize pedometers for fitness purposes. Since the majority of modern smartphones come equipped with an integrated accelerometer, we were able to employ smartphones in this project to implement pedometer functionality. This pedometer was used in our project to reduce the cost of the Fitbit devices, which we now spend between 5K-6K.

D. DATASET DESCRIPTION

Table 1 provides a summary of the information from three public sources. There are several obvious variations among them. There are the most volunteers in the UCI-HAR dataset, which indicates that person's recordings were used to create this dataset [20]. The HHAR dataset has the same six activities as the UCI-HAR dataset, but it contains more samples. Additionally, the dataset is unbalanced, as will be discussed later. The 6 activities that make up the PAMAP2

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

dataset. It was gathered by five different sensors types: magnetometers, object sensors, gyroscopes, ambient sensors, and accelerometers.

A. HHAR: The HHAR, [17] contains information from the twelve distinct gadget's accelerometers and gyroscopes —eight smartphones and four smartwatches used by different people while they engaged in 6 different activities. The HHAR dataset, which consists of six activities, is a collection of 3D (x, y, and z) raw signals that were taken from a subject's waist-mounted smartphone's accelerometer and gyroscope [6]. Thirty participants between the ages of 19 and 48 were included in the trials. Each individual engaged in six different activities: walk, moving up or down stairs, sitting, standing, and lying down. The dataset consists of 7,462 train samples and 2,967 test samples, respectively.

B. PAMAP2: Data from 6 different physical activities are included in the Physical Activity Monitoring dataset. We only took into account information from the inertial measuring units (IMU), which were placed in three separate body parts during the measurements (hand, chest, ankle). A total of 98209 samples makes up the PAMAP2 dataset, there are 9 subjects in the paper. while carrying an Android phone in their front leg pockets, these participants went about their daily routines. With a sample frequency of 20 Hz, the sensor in use is an accelerometer. Another component of the smartphone is a made motion sensor. Standing, sitting, walking, upstairs or down stairs, and laying were the six actions that were recorded. To guarantee the accuracy of the data, a committed individual monitored the data collecting. With the objective of illustrating the features of each axis raw data, displays the acceleration waveform of each activity for a total of 2.56 seconds (128 points).

C. USC-HAD: The Dataset makes use of highly accurate specialized hardware, focuses on the diversity of topics, and balances the person's based on gender, age, height, and weight. 14 individuals recorded between the ages of 19 and 48, the UCI- HAR dataset [30] was created. All participants in the recording were given instructions on how to conduct themselves. Additionally, they sported a smartphone around their waist that was equipped with inertial sensors. The six daily actions include walking (Walk), lying down (Lay), walking upstairs (Up), and walking downstairs (Down). This dataset also contains positional transitions from standing to sitting, sitting to standing, standing to laying, and sitting to lying, lying to sitting standing to standing, which happen between the static postures. Because postural shifts make up a small fraction of all activities, only six fundamental ones were c hosen for this paper's input samples. For the purpose of manually labelling the data, the studies had been recorded. The researchers then recorded data on 3-axial angular velocity and acceleration at a constant 50 Hz rate. Statistics show that there are 7352 samples in this collection.

TABLE 1. Dataset.

Datasets	Subjects	Activities
HHAR	9	6
PAMAP2	9	6

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

USC-HAD	14	6

E. DATA PRE-PROCESSING

It's necessary to pre-process the raw data that motion sensors collect in the following ways in order to provide a specific data dimension to the suggested networks and improve the model's accuracy. The sensors worn by the individuals are wireless, and the datasets indicated above are accurate. As a result, During the gathering procedure, some data might be lost. When this happens, NaN/0 is typically used to indicate the missing data. In order to solve this issue, the missing values in this study were filled using the linear interpolation approach.

The input data must be normalized to the 0 to 1 range since training models with large values from channels directly might result in training bias. A whole human activity recognition model was put into effect in this work. The model receives a data sequence as its input. Short time series were taken from the initial sensor data to create the sequence. The data were continuously recorded during the process of gathering data. The data gathered by motion sensors was segmented with a 50% overlap rate on sliding window in order to maintain the temporal link in an action between the data points. The sliding window's length for the datasets is 128.The recordings of each activity in the dataset are brief, requiring the use of a brief sliding window to split the data and produce more samples. It is essential to remember that we chose the optimal window size using an empirical and adaptive approach to achieve large segments for all the activities.

F. MODEL

In this paper, A 2-layer CNN-LSTM hybrid LSTM network is suggested for enhancing recognition performance. Two convolutional layers and Two LSTM layer make up the 2D CNN-LSTM. Convolutional neural networks (CNN), a subset of deep learning networks, are suggested for enhancing problem-solving capabilities in wearable-based HAR. A particular class of DL network called CNNs is capable of autonomously extracting spatial characteristics from unprocessed sensor input [24]. The need for time- series data arises from a variety of human activities, which causes the temporal interdependence. Long Short- Term Memory (LSTM) networks have been suggested as a solution to this temporal dependency problem, and their use in HAR is currently expanding. By merging many prior LSTM layers that extract the temporal features in combination with CNN layers that extract the spatial features, hybrid LSTMs give the benefits of both LSTMs and CNNs.

This limitation could be removed by using LSTM, a kind of RNN [25]. Due to its unique memory cells, LSTM has a significant advantage over convolutional neural networks in terms of feature extraction from sequence data. To more effectively extract the temporal characteristics from the sequence information in this study, the input data initially passes through two LSTM layers. 32 memory cells create an LSTM layer. Each memory cell's activity is controlled by the inputs being delivered to various gates, such as input, forgetting, and output

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

gates. At its core, machine learning is concerned with the prediction and identification of patterns, as well as the generation of appropriate outputs based on such information. Asa result of their ability to analyze data in search of previously unknown patterns, deep learning algorithms may acquire new knowledge. To improve with each new attempt, a DL model will learn.

We use the deep learning method to identify and categorize activities. The model accepts nine signals; in order to conduct a controlled experiment and for experiment reasons, we choose a 2D convolution layer with kernel size = 3, which establishes the convolution window's size, and 64 filters. Using ReLU activation, this layer. The other parameters were kept at their default values. The feature data from this part will be formatted by the flatten layer so that it can be used by the LSTM layer in the following phase.

Next, a flatten layer and a 2D maxpooling layer [26] with a pool size of 2 were added. The data that the convolution layer uses is not the same as the input that the LSTM layer accepts due to the nature of the convolution layer. We also need to use a technique to solve this problem because the data we are working with is temporal. The Time Distributed wrapper offered by Keras maintains the temporal integrity of the LSTM layer(s) while applying convolutions to the input signal and accepting a layer as an argument.

Then, an LSTM layer with 128 units and a ReLU activation is used as input to flatten the feature maps from the preceding layers. The LSTM layer or layers extract the signal's temporal dependencies. The best models for handling signal data, which is sequential in nature, fall within the recurrent neural network (RNN) group, which the LSTM network is a member of. The use of the LSTM network has a number of benefits over other deep neural networks [27]. With the exception of the number of units, which we adjust to 128, all of the layer's default parameters are left alone.

Before a model's efficacy can be assessed, the available data must be split into training and test sets. First, we divide the data into a training set (consisting of 70%) and a testing set (consisting of 30%) to ensure that our models are well-trained. Our model's accuracy was further enhanced by the addition of many metrics for gauging its results. In this case study, we examined potential indicators of a borrower defaulting on a loan. It's not simply accuracy that matters when evaluating a model's efficacy; metrics like the confusion matrix and accuracy should be looked at as well.

For this paper, data is gathered from several users as they performed in common daily actions including walking, sitting, standing, lying down, and climbing and descending stairs for a particular amount of period.

In every instance, data is gathered at a rate of 20 samples rate per second, or one record every 50 millisecond. Six columns make up the dataset: "user," "activity," "timestamp," "x-axis," "y-axis," and "z-axis." The words "user" and "timestamp" refer to the user ID and the Unix timestamp, respectively, while the remaining characters represent the accelerometer measures

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

acceleration along the x, y, and z axes/dimensions at a specific point in time. Activity is our goal variable (class-label), which we want to predict. The smart device sensor data may be used to classify the activity carried out by the smartphone user to the suggested Hybrid LSTM-based HAR framework in this paper. The overall approach taken in this paper to accomplish the research goal. The proposed hybrid LSTM-based HAR is offered to improve the LSTM- based DL networks' recognition effectiveness. During the first stage, the raw sensor data is divided into two main subsets: raw training data and test data. The raw training data is subsequently divided into 25% for model validation and 75% for training in the second stage of model training and hyperparameter tweaking [28]. The validation data are used to test five hybrid LSTM-based models, and the trained models' hyperparameters are then tuned using a Bayesian optimization strategy. The recognition performance of the hyperparameter-tuned models will next be compared with the test results.

As far as we know, the hyper parameters we selected optimum and had minimal effect on how effectively the related models performed, despite the fact that some of them were constant throughout all models. This is due to the fact that numerous, repeated experiments with different hyperparameters slightly changed the performance of the provided models. These hyperparameters included a mini-batch size of 32, 15 training epochs, and a learning rate, lr, which was set at 0.0050.

IV. EXPERIMENT RESULTS

In this paper, The LSTM-2D CNN model's generalizability and accuracy were assessed using three popular public datasets. They were all constantly recorded, and a typical technique is to divide the sensor data utilizing a fixed-length sliding window. In order to more accurately assess the model performance, a subset of the dataset was utilized to create the test set, which is completely segregated from the training set.

Pred / True	LAYING	SITTING	STANDING	WALKING	WALKING_DOWNSTAIRS \
LAYING	537	0	0	0	0
SITTING	7	386	98	0	0
STANDING	0	61	471	0	0
WALKING	0	0	2	462	28
WALKING_DOWNSTAIRS	0	0	0	0	419
WALKING_UPSTAIRS	0	2	2	1	32

Pred / True	WALKING_UPSTAIRS
LAYING	0
SITTING	0
STANDING	0
WALKING	4
WALKING_DOWNSTAIRS	1
WALKING_UPSTAIRS	434

Figure 5. Confusion Matrix

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

The database for the dataset was created using the recordings of 6 subjects carrying out 6 activities. recordings were used to create the training set, while the remaining ones were used to create the test set. Using the validation set, we train our model for 15 epochs while monitoring accuracy and error. Cross-entropy loss against several epochs during training and validation [22]. The hybrid LSTM model appears to learn well, as shown by accuracy higher than 91% and cross-entropy loss significantly less than 0.4 for both validation and training information.

With a cross-entropy loss of 0.04, the trained model performed well on the test dataset, achieving over 91% accuracy. As shown the Figure 5, The Confusion matrix shows that the two most popular activities in our sample, laying and standing, are correctly identified with a high degree of accuracy. Even though sitting and walking downstairs are minority classes, our model can distinguish between them with accuracy. For activities taking place walking upstairs and walking, accuracy is not as good as it is in the other classes. This is to be expected since the underlying information might not be enough to distinguish between these two actions since they are so similar.

Figure 6, The graph represents the count and activity as x-axis and y-axis that means the number of datapoints per activity for HAR. Figure 7, Represents the predicted values and actual values for HAR. The accuracy results for hybrid deep learning models and we considered, on the three datasets. This model shows an accuracy with an average, 6% higher than the previous

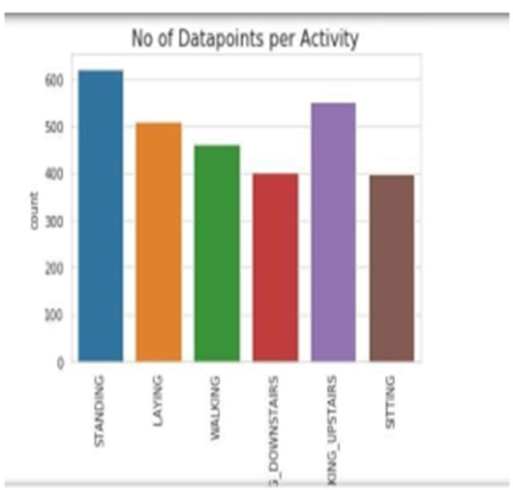
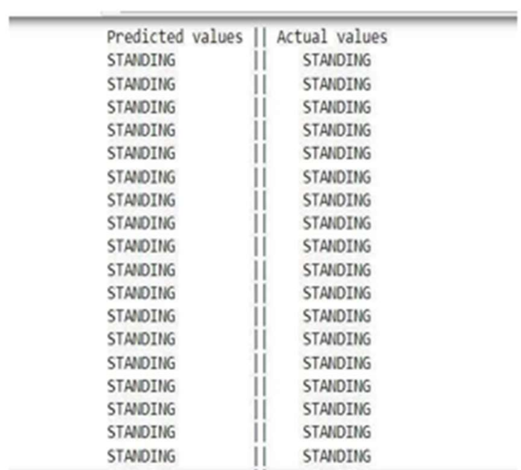


Figure 6. Results for the Models



Predicted values	Actual values
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING
STANDING	STANDING

Figure 7. Results for the Models

TABLE 2. Comparison of LSTM-2D CNN and Other Models.

System	Accuracy (%)
LSTM-2D CNN	**91**
TRASEND	**85**

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

TRASEND-CA	**79.7**

V. CONCLUSION AND FUTURE WORK

Recognition of human activity is widely used in human survey systems and medical research. In this project, we created a recognition system based on smart gadgets. Through this paper, we can identify six tasks in total (standing, sitting, walking, moving upstairs, moving downstairs and laying). The system used an integrated accelerometer to gather time series signals, produce features in both the time and frequency domains, and choose the features to increase efficiency. In order to improve the accuracy of activity identification, we suggested a 2D CNN-LSTM model for human activity detection that takes use of a CNN network's robustness in feature extraction while employing an LSTM model's work for time series prediction and classification. When compared to other deep learning techniques that employ raw signal data as input, our 2D CNN-LSTM model performed better since it is both spatially and temporally deep learning. The model was trained and tested using the activity data.

In this paper, the hybrid deep learning model achieved the greatest accuracy, 91%. This type is reliable for both smartphone orientation and placement. By taking into account predicted accuracy, these LSTM- 2D CNN networks were assessed using the HAR dataset, which is freely accessible. Our results show that the average increase is greater than 6%. We also show our model's efficacy.

Future work will take into account more activities and produce a real-time smartphone app. Additionally, we can add additional features to our model to increase accuracy, and we can use density- weighted algorithms and variance reduction to address various query tactics.

REFERENCES

[1] Y. Asim, M. A. Azam, M. Ehatisham-ul-Haq, U. Naeem, and A. Khalid, "Context-aware human activity recognition (CAHAR) in-the Wild using smartphone accelerometer," IEEE Sensors J., vol. 20, no. 8, pp. 4361–4371, Apr. 2020. doi: 10.1109/JSEN.2020.2964278.

[2] D. Chen, S. Yongchareon, E. M.-K. Lai, J. Yu, and Q. Z. Sheng, ''Hybrid fuzzy c-means CPD-based segmentation for improving sensor-based multiresident activity recognition,'' IEEE Internet Things J., vol. 8, no. 14, pp. 11193–11207, Jul. 2021, doi: 10.1109/JIOT.2021.3051574.

[3] K. Bouchard, J. Hao, B. Bouchard, S. Gaboury, M. T. Moutacalli, C. Gouin- Vallerand, H. Kenfack, Ngankam, H. Pigot, and S. Giroux, ''The cornerstones of smart home research for healthcare,'' Smart Innov. Syst. Technol., vol. 93, pp. 185–200, Apr. 2018.

[4] T. Zebin, P. J. Scully, and K. B. Ozanyan, ''Human activity recognition with inertial sensors using a deep learning approach,'' in Proc. IEEE SENSORS, vol. 1, Oct. 2016, pp. 1–3. doi: 10.1109/ICSENS.2016.7808590.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

[5] N. Ahmed, Rafiq, and Islam, “Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model,” Sensors, vol. 20, p. 317, 01 2020.

[6] C. Shiranthika, N. Premakumara, H. -L. Chiu, H. Samani, C. Shyalika and C. -Y. Yang, "Human Activity Recognition Using CNN & LSTM," 2020 5th International Conference on Information Technology Research (ICITR), 2020, pp. 1-6, doi: 10.1109/ICITR51448.2020.9310792.

[7] C. Shiranthika, N. Premakumara, H. -L. Chiu, H. Samani, C. Shyalika and C. -Y. Yang, "Human Activity Recognition Using CNN & LSTM," 2020 5th International Conference on Information Technology Research (ICITR), 2020, pp. 1-6, doi: 10.1109/ICITR51448.2020.9310792.

[8] A. Jain, K. Gandhi, D. K. Ginoria and P. Karthikeyan, "Human Activity Recognition with Videos Using Deep Learning," 2021 International Conference on Forensics, Analytics, Big Data, Security (FABS), 2021, pp. 1-5, doi: 10.1109/FABS52071.2021.9702599.

[9] A. Jain and V. Kanhangad, ‘‘Human activity classification in smartphones using accelerometer and gyroscope sensors,’’ IEEE Sensors J., vol. 18, no. 3, pp. 1169–1177, Feb. 2018.

[10] N. Jalloul, F. Poree, G. Viardot, P. L’Hostis, and G. Carrault, ‘‘Activity recognition using complex network analysis,’’ IEEE J. Biomed. Health Informat., vol. 22, no. 4, pp. 989–1000, Jul. 2018.

[11] K. Ashwini, R. Amutha and S. Aswin raj, "Skeletal Data based Activity Recognition System," 2020 International Conference on Communication and Signal Processing (ICCSP), 2020, pp. 444-447, doi: 10.1109/ICCSP48568.2020.9182132.

[12] C. Wei, H. Zhang, L. Ye and F. Meng, "A school bullying detecting algorithm based on motion recognition and speech emotion recognition," 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), 2020, pp. 276-279, doi: 10.1109/ICHCI51889.2020.00066.

[13] A. E. Minarno, W. A. Kusuma, H. Wibowo, D. R. Akbi and N. Jawas, "Single Triaxial Accelerometer-Gyroscope Classification for Human Activity Recognition," 2020 8th International Conference on Information and Communication Technology (ICoICT), 2020, pp. 1-5, doi: 10.1109/ICoICT49345.2020.9166329.

[14] P. Agarwal and M. Alam, ‘‘A lightweight deep learning model for human activity recognition on edge devices,’’ 2019, arXiv:1909.12917. [Online]. Available: https://arxiv.org/abs/1909.12917.

[15] C. Xu, D. Chai, J. He, X. Zhang, and S. Duan, ‘‘InnoHAR: A deep neural network for complex human activity recognition,’’ IEEE Access, vol. 7, pp. 9893–9902, 2019.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

[16] S. Joshi and E. Abdelfattah, "Deep Neural Networks for Time Series Classification in Human Activity Recognition," 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2021, pp. 0559-0566, doi: 10.1109/IEMCON53756.2021.9623228.

[17] D. Buffelli and F. Vandin, "Attention-Based Deep Learning Framework for Human Activity Recognition With User Adaptation," in IEEE Sensors Journal, vol. 21, no. 12, pp. 13474-13483, 15 June15, 2021, doi: 10.1109/JSEN.2021.3067690.

[18] W. A. Kusuma, A. E. Minarno and M. S. Wibowo, "Triaxial accelerometer-based human activity recognition using 1D convolution neural network," 2020 International Workshop on Big Data and Information (IWBIS), 2020, pp. 53-58, doi: 10.1109/IWBIS50925.2020.9255581.

[19] A. E. Minarno, W. A. Kusuma, H. Wibowo, D. R. Akbi, and N. Jawas, "Single Triaxial Accelerometer-Gyroscope Classification for Human Activity Recognition," 2020 8th Int. Conf. Inf. Commun. Technol. ICoICT 2020, 2020.

[20] Sansano, E., Montoliu, R., & Belmonte Fernández, O. (2020). A study of deep neural net-works for human activity recognition. Computational Intelligence, 36(3), 1113–1139. 10.1111/coin.12318.

[21] Saurabh Gupta," Deep learning based human activity recognition (HAR) using wearable sensor data," 2021, https://doi.org/10.1016/j.jjimei.2021.100046.

[22] R. Gómez, "Understanding Categorical Cross-Entropy Loss, Binary Cross-Entropy Loss, Softmax Loss, Logistic Loss, Focal Loss and all those confusing names," Online, Available https://gombru.github.io/2018/05/23/cross_entropy_loss/, Accessed: Dec. 1, 2019.

[23] M. Adjeisah, G. Liu, D. O. Nyabuga and R. N. Nortey, "Multi-Sensor Information Fusion and Machine Learning for High Accuracy Rate of Mechanical Pedometer in Human Activity Recognition," 2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), 2019, pp. 1064-1070, doi:10.1109/ISPA-BDCloud- SustainCom-SocialCom48970.2019.0015.

[24] J. Li, Y. Guo and Y. Qi, "Using Neural Networks For Indoor Human Activity Recognition with Spatial Location Information," 2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2019, pp. 146-149, doi: 10.1109/IHMSC.2019.10130.

[25] S. Bhattacharjee, S. Kishore and A. Swetapadma, "A Comparative Study of Supervised Learning Techniques for Human Activity Monitoring Using Smart Sensors," 2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC), 2018, pp. 1-4, doi: 10.1109/ICAECC.2018.8479436.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

[26] S. Yu and L. Qin, "Human Activity Recognition with Smartphone Inertial Sensors Using Bidir-LSTM Networks," 2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE), 2018, pp. 219-224, doi: 10.1109/ICMCCE.2018.00052.

[27] R. Mutegeki and D. S. Han, "A CNN-LSTM Approach to Human Activity Recognition," 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 2020, pp. 362-366, doi: 10.1109/ICAIIC48513.2020.9065078.

[28] N. Tüfek and O. Özkaya, "A Comparative Research on Human Activity Recognition Using Deep Learning," 2019 27th Signal Processing and Communications Applications Conference (SIU), 2019, pp. 1-4, doi: 10.1109/SIU.2019.8806395.

.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/