Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 212 (2022) 64–73

Procedia Computer Science

www.elsevier.com/locate/procedia

11th International Young Scientist Conference on Computational Science

FLAAP: An Open Human Activity Recognition (HAR) Dataset for Learning and Finding the Associated Activity Patterns

Prabhat Kumar* and S. Suresh

Department of Computer Science, Institute of Science, Banaras Hindu University, Varanasi – 221 005, India.

Abstract

A significant quantity of research work has been completed to recognize human activities. The majority of the proposed learning algorithms have treated the activity data as the role of fuel in vehicles. The capacity of the learning algorithms to recognize the activity patterns can be improved by the adequate availability of activity data. In this paper, we introduced the FLAAP (Finding and Learning the Associated Activity Patterns) activity dataset and data acquired by using the smartphone (accelerometer and gyroscope) sensors placed at the waist of the subjects while performing the activities. This dataset contains the record of ten activities performed by eight distinct subjects. Between February 1st and May 31st, 2022, millions of raw sensor activity data samples were captured constantly at 100Hz sampling rates. The Human Activity Recognition (HAR) datasets, which keep a record of such activities and report associations in activity patterns (mostly used for recognizing the Activities of Daily Living (ADL)), were lacking. This paucity is addressed by the FLAAP dataset which can be useful in finding the associated patterns in ADL. The obtained experimental findings demonstrate that the learning algorithm Random Forest (RF), which was used, has recognized the activities with around 77.22% accuracy. The applied RF learning algorithm on the FLAAP dataset provides the research gap for the researchers in developing more delicate learning models for enhancing recognition rates. Furthermore, the research community could be particularly interested in examining the learning performance of algorithms while using various data pre-processing techniques, transferring knowledge to target domains, and other techniques.

Keywords: Activities of Daily Living (ADL); Time Series Data; Associated Activity Patterns; Human Activity Recognition (HAR); Smartphone Sensors.

* Corresponding author. Tel.: +91-7049271390.
E-mail address: prabhat.kumar13@bhu.ac.in

10.1016/j.procs.2022.10.208

1. Introduction

In [1], the authors have discussed that there are already over 3.5 billion smartphone users worldwide, and within a few years, that figure is predicted to increase to 4 billion. In such a short period, no other technology in history has amassed among users of all ages. Even though they carry out complicated operations on the inside, smartphones are rather simple to use. It has powerful computational capabilities, and multi-functionality and is equipped with numerous sensors, and more. Despite its ostensibly simple user interfaces, the majority of installed applications nowadays are dynamic and employ several intelligent algorithms. Data gathered from user inputs and a variety of internal sensors are crucial to the learning process of smartphones. The continual collection of numerical, signal, image, and video data by using the contemporary smartphone is unrestricted until specifically told otherwise. To comprehend the essential elements of the immediate environment, these data are subsequently utilized. Additionally, noise is frequently present in the raw data. The pre-processing technique must be employed to handle the noisy data. After pre-processing the sensory data, sophisticated learning algorithms are employed to recognize the activity patterns. Through this process of learning, smartphones can more clearly comprehend the demands of the users. A prime illustration of how adaptable modern smartphones are used in Human Activity Recognition (HAR). The data from two specific sensors, the accelerometer, and gyroscope, are often taken into consideration for sensor-based HAR. An accelerometer sensor measures the linear acceleration of movements along its three X, Y, and Z axes whereas the gyroscope sensor counts the rate of angular rotation around certain axes. The change for these characteristics is recorded in the sampling rates (Hz). The fundamental idea behind activity recognition using smartphone sensors is that, as long as the device is attached to or being carried by a subject, the data can be used to describe the movements of the subjects. Further, the data from additional sources, such as the magnetometer and Global Positioning System (GPS), can also be used to enhance the activity recognition rates. In many sectors where intelligent machines are working that identify human movements, such as robots, prosthetics, haptics, surveillance systems, autonomous health care monitoring systems, and many more, HAR is widely used. Despite this, the basic objective, which is to accurately recognize human motions or gestures based on a single or many sensor outputs, remains the same. Sensor data acquisition and the study on proposing the recognition model are the two different groups in which the HAR research domain is divided. The first group focuses on collecting massive amounts of activity data from numerous subjects using smartphones or wearable sensors. And, the acquired data become usable for researchers after accomplishing certain pre-processing and noise-removal techniques on the raw sensory data. Studies in the second group are dedicated to proposing learning algorithms and applied on these data. We are from the first group and contribute to introducing the new HAR dataset, named FLAAP (Finding and Learning the Associated Activity Patterns). There are the following values of our FLAAP dataset, illustrated below:

- The activity raw data has been obtained using a smartphone-based accelerometer and gyroscope sensor which can be used to find and learn the associated activity patterns.
- The FLAAP dataset comprised recordings of eight different subjects executing ten distinct human activities at a sampling rate of 100Hz with the smartphone horizontally positioned at the waist of the human body. The sampling rates varied according to the complexity of activity patterns. Here, the sampling rates 100Hz enough to recognize the mentioned activities.
- The FLAAP dataset was the first publically accessible Indian version dataset, to the best of our knowledge. We used our daily smartphone sensors to collect the raw sensor data in a real-time environment. This is very economical and flexible to set up the experimental environment for collecting the activity data using sensors as compared to the images/videos based.
- This sensor dataset was gathered over a lengthy period, allowing researchers in the HAR area to use a variety of learning algorithms and their variant forms to demonstrate extraordinary recognition performance.
- In the FLAAP dataset, the trimmed data contained those sections whenever the subjects performed the activity. This data can be used by the researchers to create and develop an optimized model for recognizing activities.
- The various sensor devices were positioned at diverse body locations of the subject, like in previous HAR datasets. The scenario behind the use of a large number of sensors and their placement on various parts of the human body has been extensively examined. These devices were particularly sensitive to the activity being carried out. Finding and employing just the needed number of sensors, as well as determining which location of the body parts can

catch all samples for a certain type of activity, requires devotion and attention. This technique was more useful for proposing light-weighted and optimized learning models. This is the reason behind placing the sensing device at the waist of the human body.

The remainder of this paper is organized as follows. Section 2 summarizes the relevant works on relevant HAR datasets. The procedure for collecting activity data is described in Section 3. Section 4 provides a thorough overview of the Experimental environment and dataset. Section 5 reports on the HAR experimental findings and outcomes discussion. The conclusion and future works are discussed in Section 6.

2. Related Works

The introduced FLAAP dataset is freely downloadable, usable, modifiable, and accessible. But before detailing the introduced dataset, the overview of several well-known and most referenced HAR datasets, over the last ten years, have been utilized by many researchers to develop various activity recognition models, discussed in this section. According to [2], there is a lack of publicly available sensor-based HAR datasets which often keeps certain constraints like fixed sampling rates, few numbers of subjects, lack of association in activities, too similar nature of activity labeling, limited embedded number of sensors, and paucity of flexibility while extracting the features and selecting the sub-samples from sensor sensory data. The summarized information of the related sensor-based HAR datasets is mentioned in Table 1.

Table 1. Summarized characteristics of the publicly available sensor-based HAR datasets with our FLAAP and the symbol "#" denotes "total number of".

Dataset	Subjects (#)	Activities (#)	Sampling Rates (Hz)	Devices	Sensors	Ref.
Real World	15	8	50	Smartphone and Smartwatch	Accelerometer	[3]
UCI Heterogeneity AR	9	6	100 – 200	Smartphone and Smartwatch	Accelerometer and Gyroscope	[4]
Real Life	19	4	11 – 51	Smartphone	Accelerometer, Gyroscope, Magnetometer, and GPS	[5]
UCI HAR	30	6	50	Smartphone	Accelerometer and Gyroscope	[6]
WISDM	29	6	20	Wearable Sensors	Accelerometer	[7]
HARTH	22	9	50	Axivity AX3	Accelerometer	[8]
FLAAP	8	10	100	Smartphone	Accelerometer and Gyroscope	-

In the real-world dataset [3], 15 subjects carried the seven wearable devices embedded in different body parts and performed eight daily living activities, namely: stair up, stair down, jumping, lying, standing, sitting, running/ jogging, and walking. The sensor data was collected using a smartphone and smartwatches at a sampling rate of 50 Hz. The authors looked at the feasibility of using a single acceleration sensor to identify a wearable device's current on-body position in a real-world scenario during a variety of activities. In the UCI Heterogeneity AR dataset [4], the nine users carried eight smartphones and four smartwatches while performing six different activities: biking, sitting, standing, walking, stair up, and stair down. Each subject has clearly instructed to act for five minutes for each activity which helped to ensure the approximately equal data sample distribution and sampling range varied between 100Hz and 200Hz. The authors focused on accelerometer and gyroscope sensors and offered three main aspects of sensor heterogeneity categories: To begin with, there are sensor biases, which for some of the devices tested indicated an 8 percent stillness deviation from the sole exerted force, gravity – a bias significant enough to account for the acceleration of a fast train. Second, several of the devices evaluated experience severe sampling instabilities, particularly when they are doing other tasks that produce large loads. Finally, the disparity in nominal sampling rates among devices creates significant difficulty for activity recognition on-device models that have not yet been examined in the training phase. Proposed in [5], the device orientation independent, placement independent, and subject independent features have been introduced by the authors in the acquired dataset. A total of 19 subjects participated and performed the four scripted activities i.e. inactive, active, walking, and driving at the sampling rates of 11 – 51Hz. The activity samples were collected using an embedded accelerometer, gyroscope, magnetometer sensors, and GPS in smartphones. By recording 30 subjects performing the six activities by placing the smartphone on their waist, the authors have introduced the UCI-HAR dataset for activity recognition [6]. In the WISDM dataset [7], the authors have introduced the smartphone-based accelerometer sensors dataset for recognizing human activity. The dataset has been collected from twenty-nine subjects by performing the daily living activity walking, jogging, climbing stairs, sitting, and standing at the sampling rates of 20Hz. The multilayer perceptron learning algorithm has been applied to the resulting training data. In the HARTH dataset [8], the activity data has been collected from the twenty-two subjects about the time 90 to 120 minutes. The two three-axial accelerometer sensing devices have attached to the thigh and lower back and data was sampled at the rate of 50Hz. The performed activities are listed as sitting, lying, running, walking, cycling (sitting), standing, stairs up, stairs down, and cycling (standing). After comparatively studying the related HAR datasets, we found the research gap in terms of the associativity in activities, data diversion, and fixing the smartphone. The associations in activity patterns are widely referenced for recognizing the Activity of Daily Living (ADL). Due to data diversion, it is challenging to classify the streaming sensor data since different activity patterns for the same activity differ somewhat from one another. Additionally, the location of the sensing devices and the type of activity being carried out immediately correspond to one another. Finding the ideal location to fix the sensing device to obtain relevant samples is a difficult issue. We need to do cutting-edge HAR research to address these issues. And that was the driving force behind our decision to carry out the HAR data collection project for introducing the FLAAP dataset.

3. Methodology for FLAAP Data Collection

At Banaras Hindu University (BHU) in Varanasi, Uttar Pradesh, India, in a real environment, activity data for the FLAAP dataset were collected. There was a total of eight different subjects who participated in the data collection process. A variety of weight classes and ages are represented among the contestants. Table 2 presents the statistical properties of the subjects.

Table 2. Information about the subjects who gave the HAR data.

Attributes	Values
A total number of subjects:	08
Height range of the subjects (c.m.):	168 – 170
The average height of the subjects (c.m.):	169.12
Weight range of the subjects (k.g.):	55 – 90
The average weight of the subjects (k.g.):	66.00
The age range of the subjects (years):	24 - 37
The average age of the subjects (years):	29.75

Each subject received a briefing on the rationale, the method, and the intended use of the data collection before their involvement. The consent document was provided that listed all the guidelines, requirements, and possible dangers. Moreover, they also received a withdrawal form that outlined how to remove their activity data later on if they choose to do so. Each subject was free to refuse to do any activity on the list or to cease participating in the process entirely. Additionally, they were told to report any pain or injuries as soon as they occurred, both during and after the session. After accepting all the rules and asserting their rights, they signed the consent form and added their names, subject IDs, and dates. A subject was only added to the project database after satisfying these procedures. A small group of subjects was chosen to help with the data gathering process, and each of them received a separate briefing on what their responsibilities would be during the entire process. Despite the absence of such incidents, a first-aid kit was always available, and a health worker was trained on how to act without delay in the case of a serious injury. Throughout the whole procedure, emphasis was placed on the safety and willingness of the subject. At the outset of the procedure, the height, weight, and ages of subjects were noted, as well as any known medical issues. To track the samples acquired from their performed activities, each subject was given a unique subject ID. Then, the subjects have fitted with a smartphone that was strapped around their waist and had an accelerometer and a gyroscope sensor. An android application was installed on the smartphone that help to collect the sensor data. The model name of the smartphone was Samsung Galaxy M31s. Using a smartphone, a subject, and an ongoing data repository, Fig. 1. shows a systematic representation of the data collection process. Additionally, it stated that the smartphone should have been held with the left side down and the screen towards the subject when capturing the HAR data. The data was collected and stored in a secure cloud database, eliminating the need and limitation for local storage. The data was safely downloaded to a permanent storage device when the data gathering process was completed. To expedite the process, many groups of assistants worked at the same time, collecting data from various subjects. The entire operation, however, was meticulously observed at all times. The samples taken from each participant were meticulously recorded in a handwritten document.

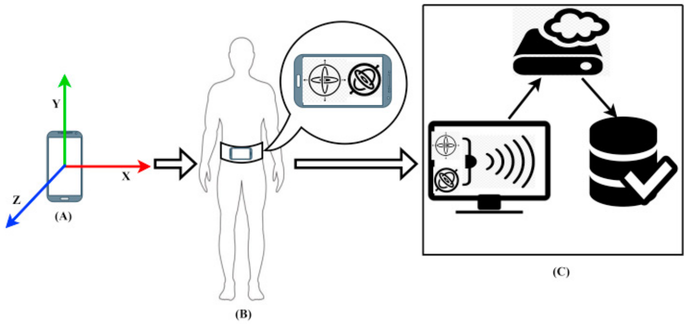

Fig. 1. A graphical depiction of the HAR data collecting process, showing (A) a smartphone for collecting the accelerometer and gyroscope sensor data, (B) the smartphone carrying belt placed at the waist of the subject, and (C) an android application installed in the smartphone that collects and fetch the sensor data to the cloud storage then downloading to the permanent storage media for further processing.

Samples from 10 distinct activity classes are included in the FLAAP dataset. Table 3 provides a description of these activities in more detail. In general, static activities and dynamic activities are broadly categorized into two groups. The data of the first four activities: sitting, standing, laying, and sitting with cross-leg activity (Activity ID 01 – 04) were designated as static type activity whereas dynamic type activity characterized by the last six activities: walking, jogging, circular walk, stair up, stair down, and sitting up (Activity ID 05 - 10). The static type activities contained less quantitative variation but the dynamic type of activities contained more quantitative variations. The working nature of the accelerometer and gyroscope sensor is defined as measuring the static or dynamic acceleration and determining the angular velocity, respectively. The static activity has been performed straightforwardly and contained fewer angular movements.

Table 3. An explanation of the conducted activity classes of the FLAAP dataset concerning average duration.

Activity	Activity ID	Activity Description	Average Duration (mm:ss)
Sitting	01	Sitting still on the garden bench	03:00
Standing	02	Standing still on the plane surface	02:48
CrossLeg	03	Sitting still on the garden bench by crossing the leg	02:57
Laying	04	Laying still on the garden bench	02:21
Walking	05	Normally paced walking	04:16
Jogging	06	Jogging straight at an average speed	04:05
Cir Walk	07	Normally paced walking along a circular path	03:59
StairUp	08	Ascending the stairs at a normal speed	02:44
StairDown	09	Descending the stairs at a normal speed	02:33
SitUp	10	Sitting on the garden seat and doing sit-ups with straight legs	02:07

The outcome sensor data for these activities contained more quantitative variation in accelerometer data as compared to the gyroscope data. Moreover, the dynamic activity has been performed in both a straightforward and angular manner and made an approximately equal effect on acceleration and angular movement. The outcome sensor data for these activities contained approximately equal quantitative variation in both accelerometer and gyroscope data.

4. Experimental environment and dataset description

This section includes the description of the experimental environment and the FLAAP dataset. The experimental environment contains a description of the subjects, a location layout, and a procedure for activity data collection whereas the dataset description explains the details of the raw data and trimmed raw data files.

4.1. Subjects

A total of eight healthy subjects participated in the data collection process. The information on the subjects has shown in Table 2. The average height (cm), weight (kg), and age (years) of the participants were 169.12, 66.00, and 29.75, respectively. All of the subjects were given a detailed explanation of the experiment they would be performing. The protocols used for collecting the activity data were explained in detail to all of the subjects. Each activity was subjected to three separate trials. After evaluating and visualizing the acquired data from those trials, we decided on the best results activity patterns. Each participant was requested to sign the consent form. This consent form has been considered a signed document that details the informed consent of the subjects for study/research on the associated patterns in their activity data.

4.2. Environmental Location Layout

The layout of the experimental location has been divided into different sub-layouts according to the nature of the activity being performed. The map view of the experimental location has shown in Fig. 2. As shown on the map, Building - 1 has located behind Building - 2. In Building - 1, the stairs are the flight of steps that leads from one floor to another, located inside the building. All of the subjects used it while performing the stairs up and down activity. For performing the remaining activities, the sub-layouts were discovered in front of Building - 2. The 65-meter long and straight path was identified front point of the department to perform the walking and jogging activity.

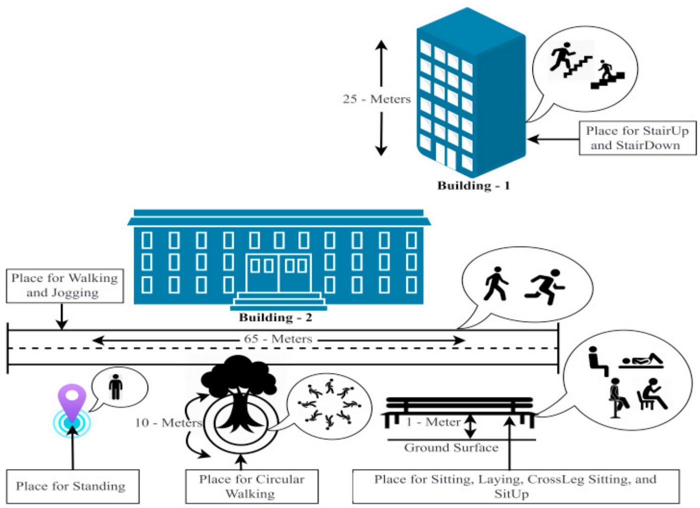

Fig. 2. The layout of the data collection environment in the experimental location.

Further, the standing activity has been performed on the spotted location point. In the front of Building - 2, a tree has been planted. We have pointed the circular path along the diameter of the same tree. The subjects performed the circular walking activity using that path. For the last four sittings, sitting with crossing legs, laying, and sitting up and down activities have been performed using the garden bench.

4.3. Procedure for Activity Data Collection

Each participant was required to perform ten different activities. However, there was no restriction to perform all the listed activities. It depended on their willingness. Before commencing the experiments, the methods necessary to complete them effectively were explained. The participants were instructed to repeat each experiment three times to obtain several occurrences of the same experiment. Ten timing scales were created to guarantee the intended experiments were carried out correctly. These timing scales show how to carry out the various experiments and what actions are involved in each of them. A succession of programmed beep sounds was also employed to alert the subject when it was time to perform the activity within the present experiment. The beep sounds were employed for the following reason:

- The commencement and conclusion of the session were indicated by the once-beep sound.
- To signify the emplacing and displacing of the smartphone, a beep sound was utilized twice.
- A beep sound was used three times to signal the start and finish of an activity.

The timing scales have all of the numbers in minutes. There was clearly described to the subjects before doing the experiments, and any questions they had about the data collecting procedure were answered. To ensure equal time across all subjects in the trial, each subject was required to follow these pre-determined timing sequences.

4.4. Dataset Description

The FLAAP dataset was grouped into one main folder that contained a total of fourteen sub-folders, arranged like a tree structure. In Fig. 3, we outlined the directory structure of the FLAAP dataset, namely – activity-wise, subject-wise, master file (.xls), raw data (.csv), raw data, and trimmed data. The activity-wise sub-folder contained the complete data for each activity that has been performed by all subjects. Similarly, the subject-wise sub-folder kept the record of all activities for each subject whatever they preferred to perform.

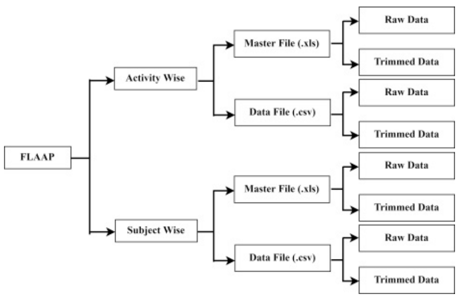

Fig. 3. The directory structure for storing the sensor data in the FLAAP dataset.

The master file contains the record in .xls format. Further, this collection includes all those columns and formulas that were required to post-processed the raw data and examine the authentic properties such as sampling rates, data consistency with continuity using a timestamp, and measuring the frequency. The outcomes of the master file were saved as the data file with limited columns (Timestamp, X, Y, Z, and Activity) in .csv format. At the end-point of the directory structure, the raw and trimmed files were found. The raw data file contained the complete activity samples without applying any post-processing techniques. So, it never ensured the quality of data because may contain noisy, missing, and outlier data.

Table 4. The fields that make up the structure of the sensor data file that was acquired are described.

Field	Description
Timestamp_Acc	When the accelerometer signal was captured at the exact moment (in ms).
Acc_X	X-axis acceleration of the accelerometer sensor (in m/s^2).
Acc_Y	Y-axis acceleration of the accelerometer sensor (in m/s^2).
Acc_Z	Z-axis acceleration of the accelerometer sensor (in m/s^2).
Activity	The name of the activity in which the subject is participating.
Timestamp_Gyr	When the gyroscope signal was captured at the exact moment (in ms).
Gyr_X	The pace at which the X-axis rotates of the gyroscope sensor (in rad/s).
Gyr_Y	The pace at which the Y-axis rotates of the gyroscope sensor (in rad/s).
Gyr_Z	The pace at which the Z-axis rotates of the gyroscope sensor (in rad/s).
Activity	The name of the activity in which the subject is participating.

However, the trimmed data file ensured the data quality because kept only those records whenever the subject started to perform an activity till its completion. The FLAAP dataset comprised both raw and trimmed data files in the activity records. The sensor data acquired while performing the activity was included in each collected data file. Furthermore, as seen in Table 4, these data files contain several fields. Two versions of the HAR data that were gathered are included in the FLAAP dataset that is available online. The accelerometer and gyroscope sensors used to capture the time-domain information from the subjects were in each of them. Each of these variations is submitted as a separate file. Here is a description of what they contain:

4.4.1. Raw data files

Without any kind of post-processing or alteration, these data files include the acquired raw activity data. Total activity samples of eight subjects, as previously noted, were gathered. Four distinct folders each contain one of these samples. The data from the tri-axial time-series sensors that correspond to the activity sample are contained in each file.

4.4.2. Trimmed raw data files

The data collection process is forced to diverge from theory and pollute the samples with noise and interference due to several causes. The interval between the beginning of the activity and the beginning of the recording is one such occurrence. Due to this discrepancy, the data collected in the first few seconds could not even be from the associated action. After the sample, the same thing can also occur. Including these elements in the sample might lead the learning algorithms astray. Therefore, we applied a trimming procedure to the impacted signals to eliminate these parts. To detect these instances and exclude them, all the gathered samples were carefully examined. The rate, however, frequently varied from 100 Hz due to variations in the capacity, processing power, and quality and design of the sensors in smartphones. To identify these instances, we recorded the precise recording time for each data point. To maintain a consistent sample rate of 100 Hz for all of the signals, we performed a one-dimensional data interpolation using the knowledge about the amount of time that had passed. The matching trimmed data files have been given together with the resultant activity samples.

5. Experimental Outcomes and Discussion

The major goal of this paper is to provide an in-depth explanation of the introduced FLAAP dataset. The activity recognition results using a classical machine learning algorithm, though, as the dataset was applied to assist the researchers to propose the novel HAR models. Various state-of-the-art HAR classification models, such as advanced deep learning and shallow learning algorithms, are now available [9]. In this study, we employed the Random Forest (RF) algorithm, a methodology for dealing with classification problems. The RF is an effective ensemble learning method that combines the results of several Decision Trees and forecasts the activity class using "bagging" concepts

[10]. For the FLAAP dataset, which was used to evaluate the performance of the learning algorithms, we randomly picked 70% of the activity samples for training. And, the remaining 30% of the activity samples were used to test the performance of the trained classifier. The RF classifier achieved 77.22% and 77.18% accuracy and F1 score, respectively. Given that the FLAAP dataset contains an uneven distribution of sample sizes for the activity classes, as shown in Fig. 4., the F1 score, which is unaffected by class imbalance, would be a better metric to evaluate the performance of the RF classifier than the accuracy rate. Fig. 5 (A) displays the confusion matrix for the RF classification that was carried out. This matrix provides more detailed information on the effectiveness of the classifier in recognizing specific activity classes. Despite having high success rates in the majority of the classes, the confusion matrix shows that the classifier has problems categorizing certain samples that are stair down, sit up, and laying behavior. This occurred as a result of how similar these three activities need different types of movement. According to recognition rates, the classes of activities for jogging, sitting, standing, and walking are divided. These activities include unique movement patterns. The classifier, however, did a mediocre job of identifying the remaining stair climbing, cross-legged sitting, and circular walking activities.

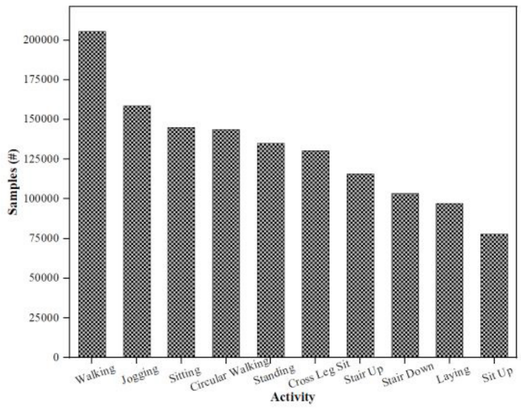

Fig. 4. Distribution of samples over the activity classes in the FLAAP dataset.

Fig. 5 (B) provides a summary of the activity class-level performance in terms of precision, recall, and F1 score. Only the stair down activity class has precision and circular walking, stair down, and walking activity classes all have rates under 70%. Finally, the experimental outcomes encouraged the researchers to critically investigate the nature of the dataset, extract distinctive and potent properties from its samples and subsamples, and develop robust activity, and recognition models.

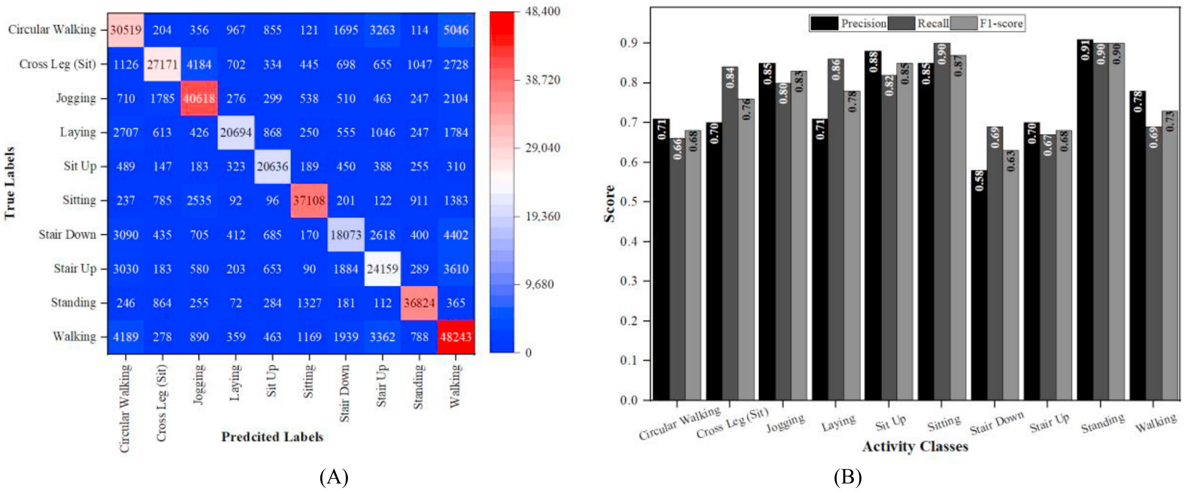

Fig. 5. (A) Confusion matrix and (B) Performance of the RF classifier of activity class-wise.

6. Conclusion and Future Work

In this paper, the introduced FLAAP dataset contained a collection of ten activity daily activities data. These data were collected using smartphone sensors. The development and formulation of the dataset were discussed in detail. Moreover, this dataset is freely usable and available online. The RF classifier results were applied to the FLAAP dataset. The experimental outcomes demonstrated the capability of the RF classifier to deliver a 77.22% and 77.18% classification accuracy and F1 score, respectively. This experimental result will be valuable for researchers. In the upcoming research work on the same dataset, the researchers will explore it in brand-new and creative ways by enhancing the recognition rates. To increase the sample sizes for certain classes and add new activities performed in daily living, we continue to collect activity data. These data will be included in the next version of the FLAAP dataset.

Data Availability Statement with Supplementary Document

The dataset presented in this article is publicly available on https://data.mendeley.com/datasets/bdng756rgw/1.

Conflicts of Interest

The authors indicate that they have no financial assistance or personal conflicts of interest that have impacted the work in this publication.

Acknowledgments

The authors like to express their gratitude to the subjects who took part in the sensor data collection studies. The authors appreciate the financial support provided by the University Grants Commission (UGC), New Delhi, through the JRF, and Banaras Hindu University, through the Institute of Eminence (IoE) Seed Grant. Moreover, the authors would like to express their gratitude to all of the faculty and researchers at the Department of Computer Science, Banaras Hindu University for their ongoing and useful discussions on this subject.

References

[1] N. Sikder, A. al Nahid. (2021) "KU-HAR: An open dataset for heterogeneous human activity recognition", Pattern Recognit Lett. **146:** 46–54. https://doi.org/10.1016/j.patrec.2021.02.024.

[2] F. Demrozi, G. Pravadelli, A. Bihorac, P. Rashidi. (2020) "Human Activity Recognition Using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey", IEEE Access. **8:** 210816–210836. https://doi.org/10.1109/ACCESS.2020.3037715.

[3] T. Sztyler, H. Stuckenschmidt. (2016) "On-body localization of wearable devices: An investigation of position-aware activity recognition", 2016 IEEE International Conference on Pervasive Computing and Communications, PerCom 2016. https://doi.org/10.1109/PERCOM.2016.7456521.

[4] A. Stisen, H. Blunck, S. Bhattacharya, T.S. Prentow, M.B. Kjærgaard, A. Dey, T. Sonne, M.M. Jensen. (2015) "Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition", SenSys 2015 - Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems. 127–140. https://doi.org/10.1145/2809695.2809718.

[5] D. Garcia-Gonzalez, D. Rivero, E. Fernandez-Blanco, M.R. Luaces. (2020) "A Public Domain Dataset for Real-Life Human Activity Recognition Using Smartphone Sensors", Sensors 2020, **20:** 2200. https://doi.org/10.3390/S20082200.

[6] D. Anguita, A. Ghio, L. Oneto, X. Parra, J.L. Reyes-Ortiz. (2013) "A Public Domain Dataset for Human Activity Recognition Using Smartphones", In: 2013 Proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (Esann),. 24–26. http://www.i6doc.com/en/livre/?GCOI=28001100131010. (accessed November 30, 2021).

[7] J.R. Kwapisz, G.M. Weiss, S.A. Moore. (2011) "Activity recognition using cell phone accelerometers", ACM SIGKDD Explorations Newsletter. **12**: 74–82. https://doi.org/10.1145/1964897.1964918.

[8] A. Logacjov, K. Bach, A. Kongsvold, P.J. Mork. (2021) "HARTH : A Human Activity Recognition Dataset for Machine Learning", 1–19.

[9] P. Kumar, S. Suresh (2021) "Deep Learning Models for Recognizing the Simple Human Activities Using Smartphone Accelerometer Sensor", IETE J Res. 1–11. https://doi.org/10.1080/03772063.2021.1967792.

[10] L. Breiman (2001) "Random Forests", Machine Learning 2001 45:1. **45**: 5–32. https://doi.org/10.1023/A:1010933404324.