SmartAberration: An annotated multimodal smartphone IMU dataset for aberration detection

Sakshi
Department of Computer Science and Engineering
Netaji Subhas University of Technology
New Delhi, India
sakshishukralia12@gmail.com
[0000-0001-9313-9308]

M.P.S Bhatia
Department of Computer Science and Engineering
Netaji Subhas University of Technology
New Delhi, India
mpsbhatia@nsut.ac.in
[0000-0001-7190-9770]

Pinaki Chakraborty
Department of Computer Science and Engineering
Netaji Subhas University of Technology
New Delhi, India
pinaki_chakraborty_163@yahoo.com
[0000-0002-2010-8022]

***Abstract*— Smartphone IMU sensors yield objective quantifications of normal and anomalous behaviors. The dataset, referred to as SmartAberration, consists of motion data collected from five participants, each engaging in 10 activities, for a total of 50 activity recordings. The data set contains both normal and aberrant activities. These are of specific interest because they have the potential to indicate security threats or interruptions of normal patterns of movement. SmartAberration dataset records motion signals derived from the smartphone IMU data, such as acceleration, velocity, and angular velocity. Each sample is labeled in terms of the actual activity performed, allowing behaviors to be classified based on their signal. This publicly accessible dataset provides a wide range of opportunities for researchers to utilize intelligent data analysis methods. It provides a platform for creating systems that can continuously monitor and evaluate normal and abnormal behaviors in real-time, without the need for self-reporting or manual observation.**

Keywords— smartphones, machine learning, soft computing, intelligent data analysis, IMU

I. Introduction

Smartphone-based sensing technology has been dramatically advancing over the past few years, offering a viable platform for collecting, analyzing, and making sense of human behavior. Among them, the Inertial Measurement Unit (IMU) sensors integrated within smartphones and smartwatches have proven themselves as robust tools for the acquisition of real-time, objective data on diverse human activities. These sensors are able to monitor various physical motions like walking, running, sitting, or standing and have also been used to monitor more intricate behaviors like sleep patterns, physical activity, and even abnormal or suspicious behavior. IMU sensors are increasingly present in consumer electronics and are a low-cost, low-hassle alternative compared to other more sophisticated and higher-cost monitoring systems [1][2].

The increasing use of smartphone-attached IMU sensors in behavioral monitoring has far-reaching implications in a range of fields, such as healthcare, security, and behavioral studies. These sensors enable round-the-clock, real-time tracking of the movements and activities of individuals without the necessity of invasive devices or wearable accessories [3].

One of the key focus areas is employing smartphones to track not only normal activities like physical activity or sleep, but even detecting abnormal or out-of-the-ordinary activities, like running in an off-limits region, assaultive behavior (e.g., punching or kicking), or other suspicious activity that could be indicative of threats or safety issues [4][5]. Having the capacity to identify and categorize these behaviors real-time and without any intervention by humans offers a critical benefit over older approaches, which are likely based on either self-report or manual observation [6].

In literature, activity recognition and behavior detection have been dominated by vision-based technologies. Security purposes have used surveillance cameras, for example, to monitor people and identify unusual behavior in public areas [7]. Vision-based techniques, however, have a number of disadvantages. Top among them is that they are extremely environment-sensitive, that is, lighting conditions, occlusion, and camera angles, which can ruin the accuracy of detecting behavior [8]. Second, vision systems tend to be privacy-invasive, with the need for cameras to record visual information of people at all times, generating major ethical concerns about surveillance and invasions of privacy [9]. In addition, vision-based technologies might find it difficult to accurately record non-visible behaviors, including subtle moves or actions carried out outside the line of the camera's gaze, hence being less trustworthy in some contexts [10].

Conversely, smartphone IMU sensors possess a number of benefits over vision-based systems. IMU sensors are not like cameras that need to record data covertly and in real-time while also violating the user's privacy. IMU sensors detect the acceleration, orientation, and angular velocity of a smartphone or smartwatch, and they provide extensive information regarding the user's movement [1][2]. Differing from cameras, which record an uninterrupted flow of visual information, IMUs concentrate on motion data alone, which can be read out in a manner that specifically measures the form, intensity, and context of body activities directly. For this reason, smartphone IMU sensors are specifically good for the monitoring of behavior in settings where visual observation is impractical, e.g., in crowded spaces, indoors, or during darkness [3]. In addition, they are not privacy invasive in the same manner that video monitoring is, and thus a more ethical and acceptable solution for around-the-clock watching.

A primary challenge in the activity recognition arena is how to separate normal from abnormal behaviors. While normal activities, like walking or jogging, are not very difficult to recognize, abnormal behaviors like running in an unusual location, sudden outbursts of aggressive actions (e.g., punching, kicking), or other abnormal actions present more challenging tasks for classification algorithms [3]. These models need to be capable of recognizing minute variations in the movement of the user, such that they can consider external context and classify behavior in a manner that closely mirrors the intent of the user or the degree of risk embodied in that behavior [10].

Therefore, this study presents a novel open dataset named SmartAberration, which to the best of our knowledge is the first to evaluate four smartphone inertial sensors (accelerometer, gyroscope, magnetometer, orientation) with GPS sensors simultaneously for the purpose of predicting aberrant activity The capacity to sense and react to anomalous behavior in real time, without manual intervention, has the potential to greatly increase personal security, safety, and general well-being monitoring across numerous environments.

II. Dataset Creation Methodology

A. *Activity classes*

This paper aims to predict and categorize human aberrations performed in restricted areas, separating routine from non-routine actions. Routine movements like walking, sitting, standing, going upstairs, and downstairs are normal and present no security risk in permitted regions. On the other hand, non-routine movements such as punching, kicking, throwing, running, and jumping, particularly along perimeter fences, are viewed as suspicious because they involve provocation or disruption. These actions are less common in normal situations and can initiate safety alarms in restricted areas.

This work employs sensor data from wrist-worn smartphones (through a wristband) and pant pocket smartphones to record accurate motion signatures. Fig.1 presents the activities logged for this research, classified as aberrant (red) or non-aberrant (green) based on their context and intensity.

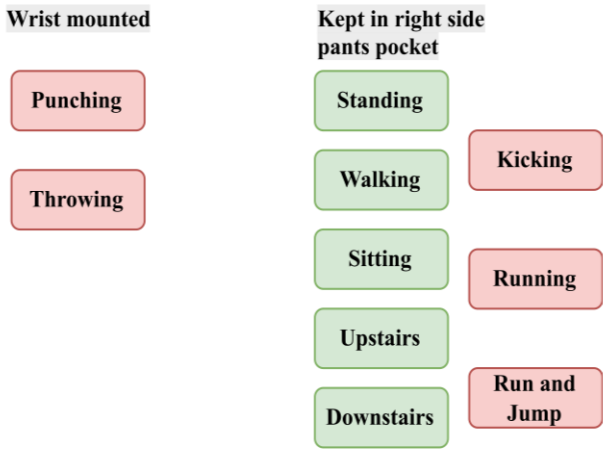

Fig. 1. Sensor location and activity classification

B. *Data collection process*

Five individuals, aged between 25 and 30 years, participated in collecting smartphone sensor data with a male-to-female sex ratio of 2:3. The participants' average height was 5 feet 4 inches (±0.215 m) and mean weight was 63.4 kg (±9.156 kg). Every participant gave written informed consent to participate and data collection. The experiment was carried out on Samsung Galaxy M31 smartphones with Android 11 installed, which contained an LSM6DSL IMU sensor. The sensor has a ±0.4% error in sensitivity and ±40mg accuracy in zero-G level offset and is capable of measuring varied things such as an acceleration and angular range.

For the removal of device sensor bias, the same smartphone model was used by all participants to ensure compatibility and dependability of data. Each participant completed 10 activities for 30 seconds, and data was captured at 50 Hz for acceleration, angular velocity, magnetic field, orientation, and position. A standardized phone placement (either within the trouser pocket or on the dominant hand) was used in an attempt to correct device placement bias. Sensor recalibration was conducted prior to every activity in an effort to reduce calibration bias. The data collection time per participant averaged 23 minutes, including 5 minutes for the activities (10 × 0.5 min) and 2 minutes of break time (9 × 2 min) between tasks.

Although most of the data were obtained within a controlled setting and under a researcher's supervision, some of these suspicious activities (e.g., kicking, punching, running and jumping, throwing) were performed following general directions but still under supervision. Going up and down stairs was conducted on staircases within the building. For ensuring a quality dataset for activities like kicking and punching, padding was applied on the walls and floors.

The IMU sensor sampled data at 50 Hz to capture different physical parameters. Although higher sampling frequencies were contemplated, it was realized that a rate of 50 Hz adequately captured the required motion features for activity prediction since higher frequencies did not improve the model's predictability. Also, the nature of the Android OS made the sampling rate provided a recommended value and not necessarily the actual rate used at times. The variation was catered for by synchronizing the data to maintain consistency in the course of the data collection [11-12].

C. Feature description

The data contains readings of the inertial sensor namely accelerometer, gyroscope, magnetometer, orientation sensors, and GPS module. All sensor streams are dated and sampled at 50 Hz, while each activity segment took 30 seconds. The following is the list of features captured in each file and their descriptions:

- X_Acceleration – Acceleration (m/s^2) along the axis aligned with the smartphone's top-bottom edges (left-right direction).
- Y_Acceleration – Acceleration (m/s^2) along the axis aligned with the smartphone's left-right edges (top-bottom direction).
- Z_Acceleration – Acceleration (m/s^2) orthogonal to the smartphone, typically pointing upward.
- X_AngularVelocity – Angular velocity (rad/s) around the axis aligned with the top-bottom edges (left-right rotation).
- Y_AngularVelocity – Angular velocity (rad/s) around the axis aligned with the left-right edges (top-bottom rotation).
- Z_AngularVelocity – Angular velocity (rad/s) around the orthogonal axis to the smartphone (vertical rotation).
- X_MagneticField – Magnetic field (μT) along the axis aligned with the smartphone's top-bottom edges.
- Y_MagneticField – Magnetic field (μT) along the axis aligned with the smartphone's left-right edges.
- Z_MagneticField – Magnetic field (μT) orthogonal to the smartphone (pointing up).
- Azimuth – Orientation angle (degrees) representing rotation around the Z-axis (compass direction).
- Pitch – Orientation angle (degrees) for tilting forward or backward around the X-axis.
- Roll – Orientation angle (degrees) for side-to-side tilting around the Y-axis.
- Latitude – GPS latitude in degrees; positive north of the equator.
- Longitude – GPS longitude in degrees; positive east of the Prime Meridian.
- Altitude – Height in meters above sea level.
- Speed – Speed in meters per second derived from GPS.
- Course – Direction of movement in degrees relative to true north.
- hacc – Horizontal accuracy in meters; radius of uncertainty around GPS position.

D. Raw data and data synchronization

We have retained raw time series data for all sensors in a different file named timetables. For every sensor data (Acceleration, Angular Velocity, Magnetic field, Orientation and Position) we have different timetables sampled at 50 Hz. Table I illustrates the used features from each sensor data and the role and direction w.r.t smartphone. Section II C describes the feature description used.

TABLE I. RAW DATA FORMAT

Sensor Type	Tuple Structure	Number of Sensors	Description
Position Sensor (GPS)	<timestamp, latitude, longitude, altitude, speed, course, hacc>	1	Captures global positioning data
Other Sensors (IMU)	<timestamp, X-axis, Y-axis, Z-axis>	4	Accelerometer, Gyroscope, Magnetometer, and Gravity sensor

In the data file, every sensor reading is put into a distinct file in the below format. Distinct activity for distinct sensor possesses distinct signatures, which are employed as features to train the model. Because of sampling difference as explained earlier, all five sensors' data were synchronized through MATLAB. Since the sensor data from all of these five sensors was gathered at the same time, hence, can be linked by their timestamps. To generate a single timetable, MATLAB function is employed that synchronizes all of the input timetables to one time vector that interpolates the values.

E. Transformed data using sliding window-data segmentation

When considering the sampling rate, it is evident that human activities typically span over extended time intervals. Consequently, a single sensor reading at a particular instant is insufficient to accurately identify an activity. Since many classification algorithms are not inherently designed to process raw time-series data, it becomes necessary to represent the data in a structured format. This is commonly achieved through data segmentation, where the continuous time-series signal is divided into smaller segments or fragments. Each segment consists of a fixed number of samples and is labeled with the corresponding activity, making it suitable for supervised learning tasks.

In this study, a time-driven segmentation approach is employed, wherein the signal is divided into successive overlapping windows of fixed duration. Specifically, the input data is segmented into 4-second windows (corresponding to 200 samples), with a 2-second overlap between consecutive windows (100 samples overlap). Each segment is then transformed into a classifier-friendly representation using a set of 18 high-level statistical and domain-specific features extracted from the raw signal. To ensure the integrity of labeled activity segments, window sizes and hop sizes are carefully chosen to avoid overlap between different activities, whether from the same or different subjects.

Additionally, window overlap is reset whenever the segmentation reaches the end of a subject's data to maintain activity consistency within each frame. From a dataset comprising 75,000 samples, this segmentation strategy resulted in the generation of approximately 700 labeled segments.

III. Experiments and Discussions

The proposed framework was evaluated on a laptop running on a Windows operating system, equipped with an i7-4510U CPU (2.0 GHz) processor and 16 GB RAM. The implementation was carried out using the Python programming language.

A. *Using deep learners- LSTM and CNN*

In this experiment, we employed a stacked LSTM model for classifying time-series data. The input shape is (200, 18), i.e., 200- time steps with 18 features per step. The network is comprised of five LSTM layers that have decreasing units (100, 75, 50, 25, 15), so that the network can learn short- and long-term dependencies of the signal. The last LSTM output is fed into a `Dense` layer of 10 units for classification. This is the most suitable architecture for sequential sensor signals such as accelerometer or gyroscope data.

In this experiment, we also employed a 1D Convolutional Neural Network (CNN) architecture for multiclass time-series data classification like accelerometer or gyroscope signals. The architecture includes a series of four Conv1D layers with increasingly and then decreasing filter numbers (64, 64, 32, 32), followed by maxpooling1d layers to down-sample the temporal dimension while preserving key features. Following the last convolution, a flatten layer converts the 3D output to a 1D feature vector of size 704. This is fed into two fully connected (dense) layers of 32 units, separated by dropout layers for regularization. The last dense layer with 10 units has a softmax activation function, because the model is being configured to classify the input into one of 10 available classes. This architecture well captures short- and long-term temporal dependencies in the sensor signal.

B. *Baseline Performance over Deep Learners*

To determine baseline classification accuracy, we used two deep learning models LSTM and CNN with the overlapping sliding window method for time-series segmentation and automatic feature extraction. Raw sensor data was divided into 4-second segments (200 samples at 50 Hz), with 50% overlap (2 seconds or 100 samples). A segment was described in terms of 18 high-level features extracted from the multivariate time series, bringing the data to the classifier input format. Both models' performance was measured in terms of accuracy and loss, providing baselines for future model development.

TABLE II. Performance over Deep Learners

Technique	Window Size	Overlap	Accuracy (%)	Loss
LSTM	4 seconds	2 s (50%)	**80.714**	0.535
CNN	4 seconds	2 s (50%)	**82.143**	0.698

IV. Applications

The SmartAberration dataset offers a rich basis for the creation of AI-powered systems that are able to continuously and non-invasively monitor human behavior in real time. These systems can be utilized across a wide range of domains to enhance safety, health, efficiency, and human activity understanding without the need for invasive manual surveillance or user input.

Applying machine learning classification models for detection and classification of normal vs. suspicious human activities. This dataset can be used in the following applications.

- Applying deep learning sequence models such as LSTM and CNN-LSTM for temporal activity recognition of IMU time-series data.
- Investigating feature extraction methods to infer statistical, temporal, and frequency-domain features from unprocessed motion signals.
- Benchmarking performances of different classifiers (e.g., SVM, Random Forest, KNN) on segmented and labelled activity fragments.
- With unsupervised learning techniques like clustering and autoencoders to find anomalies and extract latent patterns of activity.
- Employing feature selection algorithms for finding the most discriminative motion features to classify abnormal behaviors.

V. Dataset Availabilty

The data set is organized into five main folders, each of which is assigned a distinct subject (participant). Under each subject folder, there are two subfolders named Aberration and Non-Aberration, filled with data from five aberrant and five non-aberrant activity sessions, respectively. This gives a total of 50 labeled activity instances (5 subjects × 10 activities).

Each session is captured using the MATLAB Mobile App and saved in synchronized and uniformly trimmed time-series form of 30-second duration, sampled at 50 Hz. The sensor data contains readings from the accelerometer, gyroscope, and magnetometer (each with X, Y, Z axes), and GPS data (latitude, longitude, altitude, course, speed, hacc and timestamp).

For each session, there is a corresponding CSV file with the same data as well as metadata like subject ID, activity labels, and annotated timestamps. These CSV files are easily accessible, structured representations that are ideal for processing in different environments apart from MATLAB. All the files are preprocessed to maintain synchronization among all sensor modalities, with uniform time alignment and activity labeling to facilitate strong training and testing in human activity recognition and anomaly detection tasks.

The dataset can be accessed via:

https://www.kaggle.com/datasets/sakshishukralia/smartaberration

VI. CONCLUSION

This paper presents a dataset contributing in the domain of intelligent data analysis in human behavior assessing aberrations of 50 activity data for five subjects. Data is collected using smartphones inertial sensors. The dataset is open and readily available for use by researchers without requiring any preprocessing.

The dataset opens venues for the use of feature selection and extraction methods to determine significant motion patterns from raw IMU readings. These extracted features can improve machine learning classifier and deep learning model performance like LSTM and CNN-LSTM in identifying activities with high accuracy. Furthermore, unsupervised approaches like clustering and autoencoders can also be used to find anomalies and uncover hidden behavior trends. This creates prospects of designing real-time, adaptive, and intelligent systems for monitoring human activities in different fields.

REFERENCES

[1] Müller, P.N., Müller, A.J., Achenbach, P. and Göbel, S., 2024. Imu-based fitness activity recognition using cnns for time series classification. *Sensors*, *24*(3), p.742.

[2] Kim, J.H. and Lee, S.B., 2025. Evaluation of Activities of Daily Living: Current Insights and Future Horizons. *Annals of Geriatric Medicine and Research*.

[3] Kim, J.H. and Lee, S.B., 2025. Evaluation of Activities of Daily Living: Current Insights and Future Horizons. *Annals of Geriatric Medicine and Research*.

[4] Alanazi, M., Aldahr, R.S. and Ilyas, M., 2024. Human activity recognition through smartphone inertial sensors with ML approach. *Engineering, Technology & Applied Science Research*, *14*(1), pp.12780-12787.

[5] Wastupranata, L.M., Kong, S.G. and Wang, L., 2024. Deep learning for abnormal human behavior detection in surveillance videos—A survey. *Electronics*, *13*(13), p.2579.

[6] AlShami, A.K., Rabinowitz, R., Lam, K., Shleibik, Y., Mersha, M., Boult, T. and Kalita, J., 2024. SMART-vision: survey of modern action recognition techniques in vision. *Multimedia Tools and Applications*, pp.1-72.

[7] Fernandes, G.J., Zheng, J., Pedram, M., Romano, C., Shahabi, F., Rothrock, B., Cohen, T., Zhu, H., Butani, T.S., Hester, J. and Katsaggelos, A.K., 2024. HabitSense: A privacy-aware, AI-enhanced multimodal wearable platform for mHealth applications. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *8*(3), pp.1-48.

[8] Jahan, S. and Islam, M.R., 2024. A Critical Analysis on Machine Learning Techniques for Video-based Human Activity Recognition of Surveillance Systems: A Review. *arXiv preprint arXiv:2409.00731*.

[9] Dentamaro, V., Gattulli, V., Impedovo, D. and Manca, F., 2024. Human activity recognition with smartphone-integrated sensors: A survey. *Expert Systems with Applications*, *246*, p.123143.

[10] F. Cruciani et al., "A Public Domain Dataset for Human Activity Recognition in Free-Living Conditions," 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation.

[11] Lockhart, J.W., Weiss, G.M., Xue, J.C., Gallagher, S.T., Grosner, A.B. and Pulickal, T.T., 2011, August. Design considerations for the WISDM smart phone-based sensor mining architecture. In *Proceedings of the Fifth International Workshop on Knowledge Discovery from Sensor Data* (pp. 25-33).

[12] Maurer, U., Smailagic, A., Siewiorek, D.P. and Deisher, M., 2006, April. Activity recognition and monitoring using multiple sensors on different body positions. In *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)* (pp. 4-pp). IEEE.