Human Activity Recognition Using Smartphone Sensors Based on XGBoost Model

Ruikang Hu
School of Computer Science and Engineering, Southwest Minzu University, Chengdu, 610000, China

Keywords: Human Activity Recognition, Smartphone Sensors, Machine Learning, XGBoost.

Abstract: The core viewpoint of this study focuses on Human Activity Recognition (HAR) through machine learning techniques and utilizing the large amount of data brought by smartphone sensors. The increasing integration of smartphones into daily life emphasizes the need for cost-effective and convenient solutions for HAR. The goal is to explore the potential and performance of smartphone sensors in recognizing diverse activities and distinguishing between different users. This study first considers using Principal Component Analysis (PCA) as a feature dimensionality reduction and visualization analysis tool. Secondly, t-distributed Stochastic Neighbors Embedding (t-SNE) is introduced for further analysis and discussion. This paper introduces an XGBoost model for classification and contrasts it with various models. The unique feature of the XGBoost model lies in its ability to handle complex non-linear relationships, possessing high interpretability and robustness. It integrates multiple weak learners and continuously optimizes model performance through gradient boosting techniques, showcasing excellent performance in classification tasks. The experiments demonstrate high accuracy in recognizing basic activities, reaching up to 97.18%. When identifying a variety of intense sports activities, the accuracy remains high at 92.15%. In distinguishing between different users, the accuracy peaks at 93.27% for specific activities, and accurate recognition of human motion states can be achieved in less than one and a half minutes. Results highlight the feasibility of replacing traditional motion sensors with smartphone sensors, emphasizing practical applications in healthcare, fitness guidance, and gaming.

1 INTRODUCTION

Human Activity Recognition (HAR) is a technology that utilizes sensor technology and computer vision methods to monitor, analyses, and understand human movement and behavior (Kumar et al 2024, Vrigkas et al 2015 & Chen et al 2012). This technology uses various sensors such as cameras, accelerometers, gyroscopes, etc., to capture data related to human motion. Subsequently, these data are analyzed and interpreted to identify and understand various activities and behaviors of humans. With smartphones becoming indispensable companions in lives of people, their built-in multiple sensors have expanded the integration of humans with technology. In the field of HAR, the application of smartphone sensors has increasingly garnered attention (Su et al 2014). Traditional sensor-based human activity recognition methods often rely on expensive and complex wearable devices to collect data, limiting their widespread adoption in large-scale applications. However, this research focuses on using smartphone sensor data to explore the potential of a single smartphone device with machine learning techniques.

In the past few years, the field of HAR is making remarkable progress (Kumar et al 2024). Currently, researchers in the community are focusing on the application of machine learning and deep learning methods. In addition, multi-mode sensor fusion technology is also constantly developing with everyone's admiration. Secondly, the introduction of transfer learning and in-depth research on behavior modeling (LeCun et al 2015). Machine learning and deep learning models, including Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN), are pivotal in enhancing the recognition accuracy for intricate actions and activities. For example, researchers explored the use of a deep learning architecture known as Inception-ResNet for HAR (Ronald et al 2021), while Xia et. al. studied the application of the LSTM-CNN architecture (Xia et al 2020). Simultaneously, the

Hu, R.
Human Activity Recognition Using Smartphone Sensors Based on XGBoost Model.
DOI: 10.5220/0012824400004547

In *Proceedings of the 1st International Conference on Data Science and Engineering (ICDSE 2024)*, pages 286-292
ISBN: 978-989-758-690-3

fusion of data from multiple sensors, the application of transfer learning, and the ongoing demand for real-time and low-power consumption continue to drive innovation and practical applications of HAR technology. For instance, researchers successfully predicted gait freezing symptoms in Parkinson's disease using a Support Vector Machine (SVM) with a radial basis kernel (Kleanthous et al 2020). However, in-depth analysis and application of sensor data from single portable devices such as smartphones still face some challenges, including differences in data collection between different devices, the recognition capability for unconventional complex movements, and the impact of activity data on distinguishing different participants, improving recognition accuracy, and assessing activity duration. These aspects still require comprehensive and in-depth research.

The objective of this study is to explore the potential of smartphone sensors in HAR and distinguish between different users. Firstly, a comprehensive analysis of sensor data is conducted in this research, and various models are trained to validate their recognition accuracy under both routine and unconventional movements, aiming to identify the optimal model. Secondly, the study employs the XGBoost model for performance analysis in distinguishing between different participants. Experimental results demonstrate that the XGBoost model exhibits high classification accuracy, reaching up to 92.15% in complex scenarios. Additionally, there is good distinguishability among different participants. This not only highlights the reliability of smartphone sensor data but also provides practical guidance for real-world applications. It offers a solid foundation for the practical use of smartphone sensors in activity recognition. Through these steps, this paper aims to propose a more accurate, convenient, and cost-effective activity recognition solution, providing technical support for the widespread application of smartphones in areas such as human-computer interaction, anti-theft features, and gaming.

2 METHODOLOGY

2.1 Dataset Description and Preprocessing

This study primarily involves two datasets: The first dataset was collected from 30 participants engaged in daily activities such as walking, climbing stairs, descending stairs, sitting, standing, and lying down (Kaggle. 2023 a). Participants wore a waist-mounted smartphone equipped with inertial sensors (accelerometer and gyroscope) to collect data. The age range of participants in the study was set from 19 to 48 years old. Each person wears a smartphone (Samsung Galaxy S II) around their waist for six activities. The sensor of the mobile phone captures data on three-axis linear acceleration and three-axis angular velocity at a constant frequency of 50 samples per second. The dataset includes preprocessed sensor signals, initially subjected to noise filtering, followed by fixed-length interval sampling with an interval length of 2.56 seconds and 0.5 overlap (each interval contains 128 readings). The Butterworth low-pass filter is used to separate the accelerometer signal into body acceleration and gravity acceleration. Gravity, considered as low-frequency components, was filtered using a cutoff frequency of 0.3 Hertz. This dataset identifies different activities and participants separately. The second dataset records similar activities as the first, with additional activities such as cycling, playing soccer, swimming, playing tennis, jumping rope, and doing push-ups (Kaggle. 2023 b). These activities also involve body movements recorded through the smartphone's accelerometer and gyroscope sensors. Similar to the first dataset, basic data preprocessing steps were applied. This dataset only identifies different activities.

2.2 Proposed Approach

The core focus of this article is to discuss the potential of using only smartphone sensors in the HAR field. In terms of feature dimensionality reduction and visualization, Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) are used to study the label distribution of data. Subsequently, models are trained for the six fundamental activities in dataset one and the same six activities in dataset two. A comparative analysis is conducted to verify the impact of different devices on model accuracy. Additionally, training is performed on the entire dataset two, encompassing the basic six activities and an additional six sports activities, to assess whether these models can still maintain high accuracy. Finally, the research explores the distinguishability among participants in human behavior recognition, examining identification accuracy and the time required to achieve high accuracy for different participants. The process is shown in the Figure 1.

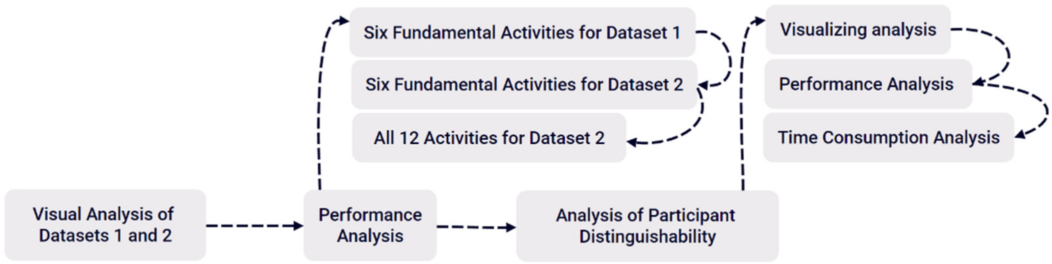

Figure 1: The pipeline of the study (Original).

2.2.1 PCA

PCA is a commonly used data dimensionality reduction technique. The technical core involves a linear transformation. Specifically, converting high-dimensional data into low dimensional data while preserving the original data information. The core principle is to identify the direction with the maximum variance in the data, known as the principal component, to achieve dimensionality reduction. PCA is characterized by decorrelation, dimensionality reduction, and maximizing variance. The main steps involve eigenvalue decomposition of the covariance matrix to determine the principal component directions and eigenvalues. In this paper, PCA is applied to process smartphone sensor data, including steps such as data standardization, covariance matrix computation, eigenvalue decomposition, selection of principal components, and data projection. This facilitates dimensionality reduction of the data, making it more accessible for analysis and understanding. The process contributes beneficial support for subsequent research endeavors.

2.2.2 t-SNE

t-SNE is a nonlinear technique used for data dimensionality reduction and visualization. t-SNE achieves this by considering the probability distribution relationships between data points, mapping high dimensional data to a low dimensional space. Minimize the Kullback Leibler (KL) divergence between two distributions to optimize the mapping. The framework of t-SNE includes computing the probability distribution of similarity between samples in the high-dimensional space, calculating the probability distribution of similarity between samples in the low-dimensional space, and minimizing the KL divergence through gradient descent to obtain the final low-dimensional representation.

In the experiments conducted in this paper, t-SNE is applied for visual analysis of smartphone sensor data. The specific implementation process first calculates the probability distribution of similarity between samples. Initialize the position of data points in low dimensional space. The algorithm optimizes the low dimensional representation through gradient descent iteration to obtain the final visualization result. In this study, t-SNE is primarily used to reveal clustering relationships between different activities, providing a more intuitive tool for data understanding and analysis in this research.

2.2.3 XGBoost

XGBoost is an efficient and powerful machine learning algorithm belonging to the Boosting type of ensemble learning. Its main features include high accuracy, strong fitting ability to complex relationships, efficiency on large-scale datasets, and support for various data types. The principle of XGBoost is based on Gradient Boosting Machine, where weak classifiers (usually decision trees) are iteratively trained. Each iteration corrects the errors of the previous round, gradually improving the overall model performance. The framework includes the optimization process of the loss function, regularization terms, and leaf node weights. Each tree is generated by minimizing the loss function, and regularization terms control the model's complexity to prevent overfitting.

The application process of XGBoost in this experiment involves model selection, parameter tuning, and training. XGBoost is chosen as the activity recognition model, and its performance metrics (such as accuracy, recall, etc.) are evaluated by comparing different models. Parameter tuning is done by adjusting key parameters like learning rate and tree depth to achieve the best model performance. The entire process includes dataset analysis, feature engineering, model training, and performance evaluation to ensure the optimal application of XGBoost in human activity recognition. The significance of using XGBoost is reflected in its successful applications in various fields, including data mining, classification, regression, and more, demonstrating its wide applicability.

2.3 Implementation Details

This research runs on the Windows operating system, and employs CPU for model training. In the system implementation, data preprocessing is carried out initially using two datasets recorded by smartphone sensors. These datasets cover the fundamental six activities as well as an additional six sports activities. To enhance the robustness of the model, data augmentation techniques are employed. Regarding the adjustment of hyperparameters, key parameters of the XGBoost model, such as learning rate and tree depth, are determined through systematic experiments. Specifically, the learning rate is configured at 0.1, and the tree depth is specified as 6. The selection of these parameters is derived through multiple rounds of experiments and cross-validation to ensure optimal performance of the model during training.

3 RESULT AND DISCUSSION

This chapter delves into the discussion and analysis of the experimental methodology employed in this paper. Firstly, PCA and t-SNE are discussed for exploratory data analysis, revealing clustering patterns of different activities in the two datasets and assessing the feasibility of classification. In terms of model training, various models are analyzed in this chapter, with XGBoost standing out. By comparing model accuracy across different devices and sensors, the superior performance of XGBoost in human activity recognition is validated. Simultaneously, utilizing XGBoost, the study confirms the distinguishability among participants and initiates an initial exploration of the time consumption aspect using smartphone sensors to explore activity habits. The following sections will provide a detailed presentation and discussion of the experimental results for visual analysis, performance analysis, and participant analysis.

3.1 Visual Analysis of Datasets 1 and 2

Figures 2 and 3 respectively depict visualizations of different activities in the first and second datasets. In Figure 2, a well-clustered pattern among the basic six activities can be observed. It is noteworthy that in Figure 3, corresponding to the second dataset, favorable clustering characteristics are still observable despite the use of different devices for collection. Additionally, another set of six more intense sports activities also exhibits relatively good classification features.

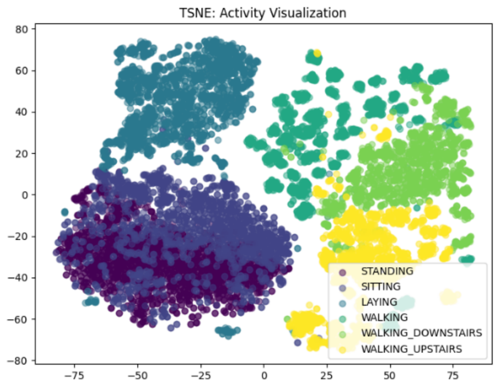

Figure 2: Activities Visualization in Dataset 1 (Original).

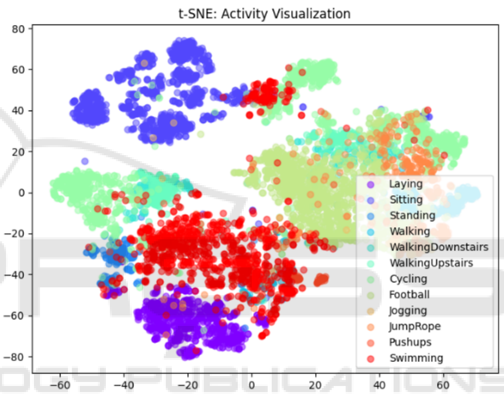

Figure 3: Activities Visualization in Dataset 2 (Original).

3.2 Performance Analysis

This study conducts model training on the six basic activities in dataset 1 and the same six basic activities in dataset 2 (results shown in Tables 1 and 2), comparing and validating the impact of different devices on model accuracy. It can be observed that the models maintain high separability under different devices. Additionally, the complete dataset 2 is trained (results shown in Table 3), including the basic six activities and an additional six sports activities. The evaluation aims to assess whether these models still maintain high accuracy. In exploring model selection, five models (logistic regression, decision tree, random forest, gradient boosting tree, and XGBoost) are chosen, and metrics such as accuracy, recall, F1 score, AUC, and precision are computed. By comparing the results in the first and second outputs, it is found that the recognition of the basic six human activities still maintains high accuracy under different devices and sensors. Even with the addition of six intense sports activities as data, the model's accuracy still remains at a high level, demonstrating the potential of smartphones in detecting intense human sports activities. It is evident that XGBoost consistently exhibits excellent performance in the aforementioned scenarios.

Table 1: Results for the Basic Six Activities in Dataset 1.

Model	Precision	Recall	F1 Score	AUC	Accuracy
Logistic Regression	0.9558	0.9557	0.9556	0.9971	0.9557
Decision Tree	0.8378	0.8377	0.8376	0.9028	0.8377
Random Forest	0.9348	0.9340	0.9335	0.9949	0.9340
Gradient Boosting	0.9419	0.9417	0.9417	0.9955	0.9417
XGBoost	0.9527	0.9522	0.9521	0.9971	0.9522

Table 2: Results for the Basic Six Activities in Dataset 2.

Model	Precision	Recall	F1 Score	AUC	Accuracy
Logistic Regression	0.9546	0.9525	0.9533	0.9950	0.9525
Decision Tree	0.9283	0.9288	0.9279	0.9591	0.9287
Random Forest	0.9761	0.9762	0.9761	0.9992	0.9762
Gradient Boosting	0.9648	0.9644	0.9641	0.9987	0.9643
XGBoost	0.9720	0.9718	0.9718	0.9989	0.9718

Table 3: Results for All Activities in Dataset 2.

Model	Precision	Recall	F1 Score	AUC	Accuracy
Logistic Regression	0.8704	0.8708	0.8681	0.9834	0.8708
Decision Tree	0.8420	0.8424	0.8413	0.9117	0.8423
Random Forest	0.9226	0.9178	0.9124	0.9934	0.9178
Gradient Boosting	0.9086	0.9073	0.9030	0.9911	0.9073
XGBoost	0.9234	0.9215	0.9166	0.9946	0.9215

3.3 Participant Analysis

3.3.1 Visualizing Analysis

Analyzing Dataset 1, in the upper part of Figure 4, a two-dimensional scatter plot illustrates the distribution of different activities, while the lower part of Figure 4 displays the two-dimensional distribution of different participants engaged in various activities. Evidently, the separability among participants is pronounced, especially in the areas related to walking up and down stairs.

3.3.2 Performance Analysis

Figure 5 depicts the model accuracy in classifying participants under different activities. As expected, the classification accuracy is higher when the physical activity intensity is elevated. This could be attributed to the sensors capturing subtle details of participants' movements, which are then identified by the model. In other words, it can analyze participants' behavioral patterns through sensor data.

3.3.3 Time Consumption Analysis

Dataset 1 involves sampling with a fixed-width sliding window, where the window width is 2.56 seconds with a 50% overlap, meaning data is collected every 1.28 seconds. As shown in Table 4, through calculations, it can be observed that in less than a minute and a half, accurate identification of human movement states (WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS) can be achieved.

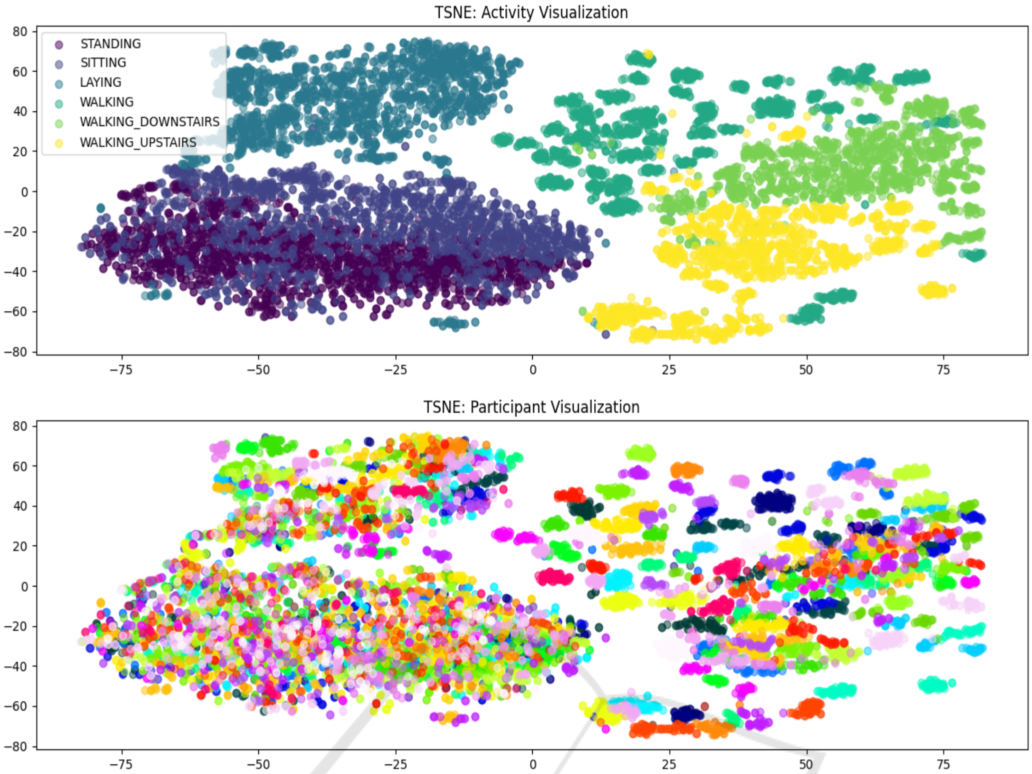

Figure 4: Activity and Participant Distribution (Original).

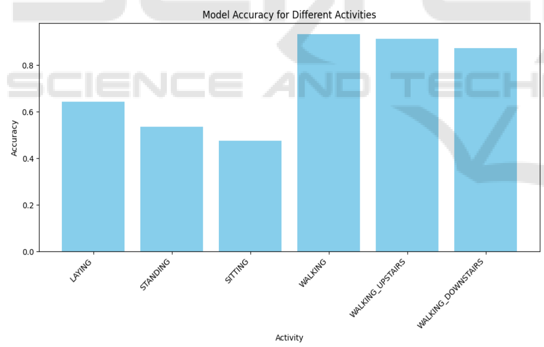

Figure 5: Participant Classification Accuracy across Activities (Original).

This chapter discusses the experimental methods and results. Firstly, exploratory data analysis was conducted using PCA and t-SNE to reveal clustering patterns of different activities in two datasets and assess the feasibility of classification. In terms of model training, an analysis of multiple models was performed, validating the outstanding performance of XGBoost in human activity recognition through comparisons of model accuracy across different devices and sensors. Under the processing of the XGBoost model, it was demonstrated that smartphones could detect user activities in less than 90 seconds. These experimental findings underscore the potential of smartphones utilizing sensors for human activity recognition.

Table 4: Identification Time Analysis.

Activity	Accuracy	Seconds
LAYING	0.641975	77.463704
STANDING	0.534591	81.754839
SITTING	0.476404	78.186667
WALKING	0.932715	73.177872
WALKING_UPSTAIRS	0.911917	66.560000
WALKING_DOWNSTAIRS	0.872159	60.416000

4 CONCLUSION

This research delves into the application of smartphone sensors in the domain of HAR. The research focuses on exploring the potential of using smartphone sensors for recognizing both conventional and unconventional activities. XGBoost, is introduced for the analysis, and extensive experiments are conducted to evaluate its performance. The model is trained and tested on two datasets, showcasing outstanding accuracy in recognizing various human activities. Through exploratory data analysis, clustering patterns for different activities and discernibility among participants are revealed, highlighting the capability of smartphone sensors to analyze human activity habits. XGBoost, proves to be effective in achieving high accuracy in HAR, even for intense sports activities. This research holds significant practical implications. Replacing traditional specialized motion sensors required for HAR with more convenient and cost-effective smartphones has improved user comfort and usability. The potential of this practical application extends beyond personal understanding of behavioral habits, providing valuable support in areas such as smartphone security, healthcare, fitness guidance, and human-computer interaction. However, the study acknowledges certain limitations. The research may not comprehensively cover the behavioral pattern variations among different participants, indicating a need for more samples and richer data to ensure reliability. Additionally, there is room for further optimization in recognizing specific complex movements, and future research will focus on enhancing accuracy and stability in this regard.

REFERENCES

P. Kumar, S. Chauhan, L. K. Awasthi, Archives of Computational Methods in Engineering (2024) pp. 179–219.

M. Vrigkas, C. Nikou, I. A. Kakadiaris, Frontiers in Robotics and AI (2015) p. 2.

L. Chen, J. Hoey, C. D. Nugent, D. J. Cook and Z. Yu, IEEE Transactions on Systems, Man, and Cybernetics (2012) pp. 790-808.

X. Su, H. Tong and P. Ji, Tsinghua Science and Technology (2014) pp. 235-249.

Y. LeCun, Y. Bengio, G. Hinton Deep learning. Nature (2015) pp. 436–444.

M. Ronald, A. Poulose, D. S. Han, IEEE Access (2021) pp. 68985–69001.

K. Xia, J. Huang, H. Wang, IEEE Access (2020) pp. 56855–56866.

N. Kleanthous, A. J. Hussain, W. Khan, P. Liatsis, Pattern Recognition Letter (2020) pp. 119–126.

Kaggle - human-activity-recognition-with-smartphones, 2023, available at https://www.kaggle.com/datasets/uciml/human-activity-recognition-with-smartphones/data.

Kaggle - inertia-sensors-for-human-activity-recognition, 2023, available at https://www.kaggle.com/datasets/owenagius/inertia-sensors-for-human-activity-recognition/data.