Proceedings of the 4th International Conference on Computing and Data Science (CONF-CDS 2022)
DOI: 10.54254/2755-2721/2/20220514

XGBoost-Based Human Activity Recognition Algorithm using Wearable Smart Devices

Ge Runchan

Harbin Institute of Technology, Harbin Heilongjiang 150001, China

1170300808@hit.edu.cn

Abstract. With the popularity of wearable smart devices, human activity recognition (HAR) based on smart sensors has been widely applied in daily life and medical health fields. In order to balance the accuracy of HAR and the complexity of the algorithm, this paper proposes an HAR algorithm based on the extreme gradient boosting (XGBoost) method. The original data collected by sensors contains noise, thus denoising process is firstly performed. Then multi-dimensional features are extracted from the data because of the limited dimensional features, which cannot be directly utilized in the training process. After that, principal component analysis (PCA) is used to reduce the dimensionality of the data in order to alleviate the input complexity of the training model. Finally, the human activity is recognized using the XGBoost method, and the ultimate goal is to obtain the tradeoff between speed and accuracy of the HAR algorithm.

Keywords: Feature extraction, human activity recognition, PCA, XGBoost.

1. Introduction

In recent years, wearable smart devices, including smart bracelets and smart watches, have been gaining great popularity. The main reason why wearable devices are so popular is that they usually have the function of recognizing and monitoring human behavioral activities, such as sitting, standing, sleeping and various sports. Most of these devices are equipped with sensors, such as accelerometers and gyroscopes to ensure the accuracy of recognition.

Human activity recognition (HAR) also has a wide range of applications in healthcare, especially for the elderly population. Healthcare service systems for the elderly using HAR with smartphone sensors, such as Mobile Health Interventions [1], have greatly improved the quality of life of the elderly [2].

Therefore, the problem of how to perform accurate recognition of human activity has attracted the attention of many researchers. 1) Traditional machine learning methods, including decision trees [2], random forests [3,4] and support vector machines [5,6], have been applied to solve the HAR problem, which are relatively simple to implement. However, the accuracy of recognition is not high enough in general because of the small number of data dimensions and the simplicity of these methods. 2) Some of the more popular deep learning methods have been introduced into the HAR problem in recent years, including Multilayer Perceptron (MLP) [7], Convolution Neural Network (CNN) [8] and Deep Belief Network (DBN) [9]. Compared with the traditional machine learning methods mentioned above, deep learning methods have achieved better results and generally higher recognition accuracy. However, these methods spend a lot of time on training models and tuning the parameters, while there is the risk

DOI: 10.54254/2755-2721/2/20220514

of overfitting, meaning that it is difficult to get an optimal model for human activity recognition. Therefore, how to balance the accuracy of recognition and the training cost of the model is an interesting research problem.

In order to solve the problem mentioned above, we propose a new method using wearable smart devices. The collected human activity data is first pre-processed and feature extracted, then PCA method is introduced to reduce the dimensionality of the data, and finally Extreme Gradient Boosting (XGBoost) algorithm is used to train the model and recognize the human activities. XGBoost is a decision tree-based algorithm, which is very suitable for the HAR problem because it converges faster and is less prone to overfitting than the methods mentioned above. In addition, the model is trained faster than any general neural network due to the support of parallelization.

2. Related works

Research on HAR with sensors has attracted many researchers, and has led to advanced results. Machine learning methods including decision trees, random forests, and support vector machine (SVM) have been applied to the recognition of human activity. Fan et al [3] used the ID3 decision tree method for HAR and compared with other two models, SVM and neural networks based on back propagation, finally concluded that decision trees have the best recognition accuracy. Guo [6] et al. focused on more specific activities, such as walking with a cell phone in hand and voice call. Then Markov chain, plain Bayesian network and SVM are implemented for comparison. Finally, it is shown that SVM had the best performance. Balli S [4] et al. applied random forest, C4.5 decision trees and SVM with data from smart watches, and concluded that random forest had the highest accuracy. Studies mentioned above achieved certain recognition accuracy on specific datasets, but the models are difficult to be widely applied to other datasets due to lack of generalization capability.

Deep learning methods have also been introduced into the field of HAR. Applying smartphone-based recognition of human activities and postural transitions data set from UCI [10], Mohammed Mehedi Hassan [9] et al used a kernel PCA approach to reduce the dimensionality of the data and later train a DBN model to recognize human activities. Hongkai Chen [8] et al. improved the traditional manual feature selection by converting 2D time series data acquired from multiple sensors into 3D tensor data and used a new CNN for HAR, which was eventually validated on the MobiAct dataset [12] with an accuracy of 99%. Using data collected from accelerometer sensors, Abdul Rehman Javed [7] et al. compared decision trees, logistic regression and multilayer perceptrons for HAR and finally concluded that MLP had the best performance and that y-axis and z-axis acceleration were more critical for HAR among the three axes of acceleration. The efficiency of these deep learning methods is obviously relatively high, some even close to 100%, but the training process of such models relies on a large amount of data and a long training period, which is difficult to implement in practice.

XGBoost is currently performing well in the healthcare field. For instance, Adeola Ogunleye et al [13] used XGBoost to achieve superb results in chronic kidney disease diagnosis: accuracy, sensitivity, and specificity of 1.000, 1.000, and 1.000, respectively. L. Torlay et al [14] used XGBoost to differentiate between patients with and without epilepsy by assessing their brain activity with the help of function MR. However, the XGBoost method has not yet been applied in the field of HAR.

3. Proposed method

For the HAR problem, how to ensure the accuracy of recognition while hoping to alleviate the cost of training the model is a topic worth exploring. To address this problem, this paper introduces the XGBoost method for HAR. It mainly includes the following processes of the algorithm: acquisition and pre-processing, dimensionality reduction and model training. Firstly, the sensor data of smart devices are collected and processed, including denoising and feature extraction. Then the dimension of data is reduced by using PCA. Ultimately, the model for recognition is trained by XGBoost method.

DOI: 10.54254/2755-2721/2/20220514

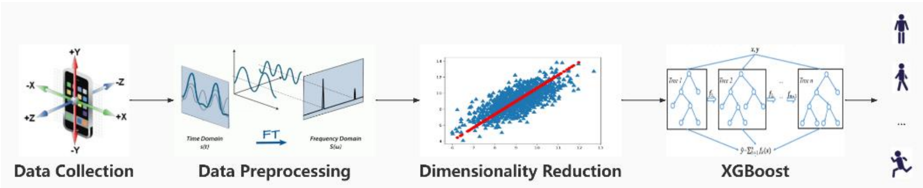

Figure 1. diagram of the XGBoost-based Human Activity Recognition Algorithm.

3.1. Algorithm description

Data Collection and Preprocessing. HAR data is generally derived from the smartphone's sensors, mainly accelerometer and gyroscope. The accelerometer sensor acquires the acceleration of human body while the gyroscope sensor acquires the direction, which is used to decompose the former into three-axis acceleration. Sliding windows are applied to the data acquisition. Assuming that the sampling frequency of the sensor is f Hz, accordingly, the step size of the sliding window is usually f. In the subsequent Fourier transform, to extract frequency domain features, the window size would be:

$$windowSize = 2^{[log_2(2*f)]} \tag{1}$$

Denoising processes are applied to the acquired raw data, using mean filters or Butterworth low-pass filters. Since the acceleration signal obtained at this point contains gravitational acceleration, further decomposition is performed to obtain the separate body acceleration. Some features can be calculated as follows:

$$S = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{t} w_{ij} \tag{2}$$

$$En = \sum_{i=1}^{N} c_i \log(c_i) \tag{3}$$

where c_i is the weight of sliding window：

$$c_i = \frac{w_i}{\sum_{j=1}^{N} w_j} \tag{4}$$

In addition, the frequency domain is sometimes well characterized by the data, and the corresponding signal in frequency domain can be efficiently calculated using the Fast Fourier Transform (FFT):

$$X(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)e^{-j\frac{2\Pi nk}{N}} \tag{5}$$

The FFT is an improvement of the discrete Fourier transform DFT. It decomposes the DFT according to the odd and even terms， and the first $\frac{N}{2}$ points and the last $\frac{N}{2}$ points after the decomposition are only different in sign, so only half of them need to be calculated. After the recursive decomposition, the complexity of the original $O(n^2)$ is reduced to $O(nlogn)$. With the frequency domain signal, some frequency domain processing can be performed, such as the spectral energy in the frequency band [a, b] range:

$$S = \frac{1}{a+b+1}\sum_{i=a}^{b} f_i^2 \tag{6}$$

Dimensionality Reduction. Multi-dimensional data generally must be dimensionalized in order to train the model in a certain time. In this paper, the PCA is utilized to reduce the dimensionality of the data. Suppose the data set from preprocessing $X = \{x_1, x_2, \dots, x_n\}$, where x_i is a certain dimension, meaning that there are $|x_i|$ samples in the data set. Our goal is to reduce the number of dimensions n to k, while maintaining the information quantity. If the dimensionality of the reduced data set $Y = \{y_1, y_2, \dots, y_k\}$, (obviously $Y = XP$), then the goal is to compute the matrix P, which is composed of k bases.

Firstly, de-mean the dataset X (X^0 obtained) and calculate the covariance matrix of X^0:

$$C = \frac{1}{n} X^0 (X^0)^T \tag{7}$$

where ${X^0}^T$ is the transpose of X^0. The obtained covariance matrix is a real symmetric matrix, and C contains all the covariances between the features.

The SVD method of eigenvalue decomposition is applied to solve for the eigenvalues $\lambda_1, \lambda_2, \dots \dots, \lambda_n$ and the corresponding eigenvectors $c_1, c_2, \dots \dots, c_n$. Then the eigenvalues are sorted and the top k largest eigenvalues $\lambda_1, \lambda_2, \dots \dots, \lambda_k$ and the corresponding eigenvectors $c_1, c_2, \dots \dots, c_k$, then the matrix $P = \{c_1, c_2, \dots\dots, c_k\}$ can be obtained. Finally, the reduced dimensional data set $Y = XP = \{y_1, y_2, \dots, y_k\}$.

Model Training. The dataset from dimensionality reduction can be used for model training and testing. Suppose the dimensionality reduction results in a dataset $X = \{x_1, x_2, \dots, x_n\}$, where x_i is a certain dimension, meaning that there are $|x_i|$ samples in the dataset.

XGBoost is essentially an additive model consisting of k base models, which are trained by continuously dividing the root node of the current tree to form a new tree with a certain strategy, and stopping the dividing under certain conditions to obtain a tree model for recognition. Suppose the tree model to be trained in the t_{th} iteration is $f_t(x)$, then

$$\hat{y}_i^t = \sum_{j=1}^{t} f_j(x_i) = \hat{y}_i^{t-1} + f_t(x_i) \tag{8}$$

where $\hat{y}_i^t$ is the prediction result for sample i after the t_{th} iteration and $\hat{y}_i^{t-1}$ is the prediction result of the first $t-1$ trees for sample i. The loss function is the deviation of the predicted value from the true value, then

$$L^t = \sum_{i=1}^{|x_i|} l\left(y_i, \hat{y}_i^t\right) = \sum_{i=1}^{|x_i|} l(y_i, \hat{y}_i^{t-1} + f_t(x_i)) \tag{9}$$

where L^t is the loss. Since only f_t is unknown, so the current goal is to learn a new function as f_t to fit the residuals of $t-1$ trees. XGBoost addresses this problem by introducing a Taylor formula for the second-order expansion of the loss function, then the loss function is required second-order derivable, such as the squared loss function:

$$l\left(y_i, \hat{y}_i^{t-1}\right) = \left(y_i - \hat{y}_i^{t-1}\right)^2 \tag{10}$$

Assume that the first-order derivative obtained after the second-order expansion is g_i and the second-order derivative is h_i, then

$$L^t = \sum_{i=1}^{|x_i|} l(y_i, \hat{y}_i^{t-1} + f_t(x_i)) = \sum_{i=1}^{|x_i|} [l(y_i, \hat{y}_i^{t-1}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] \tag{11}$$

In this case, the tree can be continuously split down by finding the first-order derivative and second-order derivative, and then optimizing the objective function to naturally obtain the model $f(x)$ for the t_{th} tree.

The accuracy of the model is determined by both the deviation, which is determined by the loss function, and the variance. the XGBoost model introduces a regularization term $\Omega(f_t)$ in the objective function to prevent overfitting. Therefore, the objective function is defined as follows:

$$Obj^t = \sum_{i=1}^{|x_i|} l(y_i, \hat{y}_i) + \sum_{i=1}^{t} \Omega(f_i) \tag{12}$$

where $\sum_{i=1}^{t} \Omega(f_i)$ is the sum of the complexity of all t trees. Additionally, the complexity of the first $t-1$ trees is less than the complexity of the t_{th} tree, so

DOI: 10.54254/2755-2721/2/20220514

$$\mathrm{Obj}^t = \sum_{i=1}^{|x_i|} l(y_i, \hat{y}_i) + \Omega(f_t) + \mathrm{constant} \tag{13}$$

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{14}$$

where T denotes the number of leaf nodes and ω_j denotes the score of the j_{th} node, while γ and λ are the parameters to be tuned during model training.

To reach an optimal model, a condition to stop splitting process is needed, since the process is recursive. The recognition capability of the model increases with the splitting the tree, so the condition should be a threshold, which indicates the least improvement of the HAR. In this case, the objective function can be as follows:

$$\mathrm{Obj} = -\frac{1}{2}\sum_{j=1}^{T} \frac{G_j}{H_j + \lambda} + \gamma T \tag{15}$$

where G_j and H_j are the sum of first and second order derivatives, respectively. Then the gain of splitting can be calculated:

$$\mathrm{Gain} = \mathrm{Obj}_{\mathrm{before}} - \mathrm{Obj}_{\mathrm{after}} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{16}$$

where $\mathrm{Obj}_{\mathrm{before}}$ indicates the value of objective function before splitting while $\mathrm{Obj}_{\mathrm{after}}$ indicates the value after the splitting. G_L G_R H_L H_R denotes the first-order derivative and second-order derivative of the left subtree and right subtree, respectively. Then γ determines the propensity to split: the smaller γ is, the higher the gain of splitting, then the more the tree tends to split. In this way, a complete XGBoost tree can be trained by tuning the value of γ.

3.2. Algorithm implementation

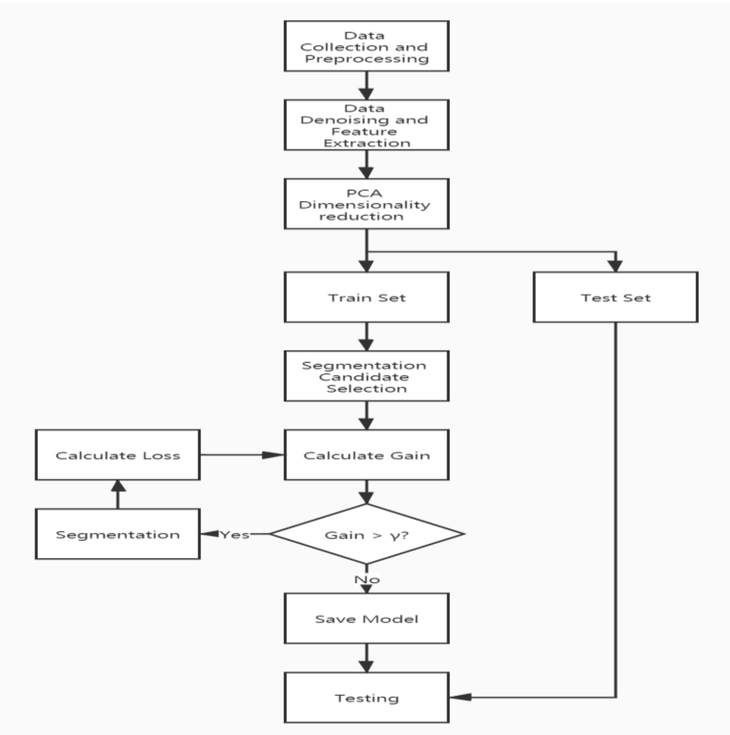

Figure 2. Process of Algorithm Implementation

1.Data collection and pre-processing: HAR data is generally collected by the sensors in smart devices, which are mainly accelerometer and gyroscope. The accelerometer sensor collects the acceleration of

the user, and the gyroscope sensor collects the direction of the user. Then the acceleration of gravity is separated from the collected acceleration data, and only the acceleration part of the body is retained.
2.Denoising process and feature selection: The data is denoised using filters. Different features are obtained by doing some basic operations on the data, such as mean, variance, sharpness, magnitude, etc.
3.PCA dimensionality reduction: the covariance matrix is firstly calculated, then eigenvalue decomposition is performed and only a certain number of features to be retained.
4.Training Model: the segmentation point is selected and corresponding gain is calculated: do the segmentation and calculate loss if the gain is greater than γ, while stop splitting if the gain is less than γ.

4. Conclusions

In this paper, we propose a new method for human activity recognition: pre-processing of the data and feature extraction, dimensionality reduction with PCA method, and finally XGBoost method is applied for recognition. Compared with previous machine learning methods or neural networks trained by deep learning method, the method proposed in this paper balances the accuracy of recognition and the complexity of the model, which can be applied in scenarios with limited resources, such as wearable devices with low performance. In this paper, we illustrate that the algorithm in this paper can effectively solve the problem of sensor-based human activity recognition through specific analysis.

In the follow-up work, we plan to use public datasets or collect data from wearable devices, such as smart watches, to verify the correctness and accuracy of the model. In fact, the data have some differences when different individuals perform the same activity, which mainly come from individual characteristics, such as gender, age, and physical condition. These individual characteristics can be used in future work to identify human activities and improve the accuracy of the model.

References

[1] Kuerbis A, Mulliken A, Muench F, et al. Older adults and mobile technology: Factors that enhance and inhibit utilization in the context of behavioral health[J]. 2017.
[2] Wang Y, Cang S, Yu H. A survey on wearable sensor modality centred human activity recognition in health care[J]. Expert Systems with Applications, 2019, 137: 167-190.
[3] Fan L, Wang Z. Human activity recognition model based on location-independent accelerometer[J]. Application Research of Computers, 2015, 32(1): 63-66
[4] Balli S, Sağbaş E A, Peker M. Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm[J]. Measurement and Control, 2019, 52(1-2): 37-45.
[5] ZHU Xiangbin, QIU Huiling. Human activity recognition with smartphone sensor data. Computer Engineering and Applications, 2016, 52（23）:1-5
[6] Guo Yuanbo, Kong Jing, Liu Chunhui, Wang Yifeng. Activity recognition based on the alteration of multi-sensor in smart phone [J].Journal on Communications,2018,39(S2):164-172.
[7] Abdul Rehman Javed, Muhammad Usman Sarwar, Suleman Khan, Celestine Iwendi, Mohit Mittal, Neeraj Kumar. Analyzing the Effectiveness and Contribution of Each Axis of Tri-Axial Accelerometer Sensor for Accurate Activity Recognition[J]. Sensors,2020,20(8):
[8] H. Chen, S. Mahfuz and F. Zulkernine, "Smart Phone Based Human Activity Recognition," 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2019, pp. 2525-2532, doi: 10.1109/BIBM47256.2019.8983009.
[9] Mohammed Mehedi Hassan,Md. Zia Uddin,Amr Mohamed,Ahmad Almogren. A robust human activity recognition system using smartphone sensors and deep learning[J]. Future Generation Computer Systems,2018,81:
[10] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21th European

Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.

[11] Chen T, Guestrin C. Xgboost: A scalable tree boosting system[C]//Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016: 785-794.

[12] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Pediaditis, and M. Tsiknakis. 2016. The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones. Proceedings of The International Conference On Information and Communication Technologies for Ageing Well And E-Health.

[13] Ogunleye A, Wang Q G. XGBoost model for chronic kidney disease diagnosis[J]. IEEE/ACM transactions on computational biology and bioinformatics, 2019, 17(6): 2131-2140.

[14] Torlay L, Perrone-Bertolotti M, Thomas E, et al. Machine learning–XGBoost analysis of language networks to classify patients with epilepsy[J]. Brain informatics, 2017, 4(3): 159-169.