ISSN 2616-5775 Vol. 6, Issue 6: 115-120, DOI: 10.25236/AJCIS.2023.060618

Human behavior recognition based on machine learning

Jingkai Zhang[1,*], Tiandao Luo[2]

[1]*School of Information Engineering, Zhengzhou Technology and Business University, Zhengzhou, 451400, China*
[2]*School of Electrical and Control Engineering, Henan Urban Construction University, Pingdingshan, 467000, China*
[*]*Corresponding author: 1965631327@qq.com*

Abstract: *At present, human action recognition technology has been widely used in intelligent monitoring, fatigue driving warning, fall detection, family rehabilitation training and other fields. In order to accurately identify human actions, this paper uses a variety of machine learning models. At the same time, in order to improve the accuracy of recognition, this paper uses a variety of feature data sets to train the model. Through experiments, it is found that the model trained by the feature data set after PCA dimensionality reduction has the best comprehensive effect. The prediction accuracy of logistic regression algorithm, KNN algorithm and LightGBM algorithm has been significantly improved. Compared with the models trained by other feature data sets, the recognition accuracy has been improved by 6% -20%, reaching 0.89, 0.87 and 0.83 respectively.*

Keywords: *Human Behavior Recognition, Machine Learning, PCA, Grid Search, LightGBM*

1. Introduction

There are many ways to extract human action signals, which are divided into two categories: image-based action signal extraction and sensor-based action signal extraction. Image-based signal extraction: By extracting the information in the picture, the key points of the human body are used as pixel coordinates. It is very convenient to use the image motion signal extraction in the scene with monitoring equipment, but the lighting conditions are higher. Sensor-based motion signal extraction: Wearable sensor devices are usually used to collect data. The commonly used inertial sensors include accelerometers, gyroscopes, and magnetometers. It is easy to extract data by wearing micro-inertial sensors and the price is low, but each extraction information requires multiple sensors to be installed in different parts, which is troublesome for the subject.

In this paper, the data obtained by the inertial sensor[1] is used. According to the extracted data, the characteristics of 23 kinds of action signals are extracted in the time domain and frequency domain, and then classified by various machine learning methods[2], so as to achieve the purpose of studying the advantages and disadvantages of each model. The three-axis inertial sensor is shown in Figure 1.

Figure 1: Three-axis inertial sensor

2. Data source

The data set used in this paper comes from 8 subjects. The subjects were equipped with 9 sensors in 5 parts of the body, and they were asked to complete 19 actions while wearing sensors: seat(A1); standing(A2); supine(A3); left lying(A4); upstairs(A5); down stairs(A6); stand in the elevator(A7);

ISSN 2616-5775 Vol. 6, Issue 6: 115-120, DOI: 10.25236/AJCIS.2023.060618

walking in the elevator(A8); walk in the parking lot(A9); on the treadmill at a speed of 4 km / h in a straight position 15 degrees inclined position(A10); walking at a speed of 4 km / h on a treadmill at a 15 degree tilt(A11); running at 8 km / h on a treadmill(A12); exercise on the stepper(A13); exercise on a cross trainer(A14); riding a exercise bike in a horizontal position(A15); ride the exercise bike in a vertical position(A16); rowing(A17); jump(A18); playing basketball(A19).

Among them, each subject is required to act for 5 minutes. Due to the different personal behavior habits of the subjects, the data extracted from the same action will be different, which increases the diversity of behavior information. The sensor unit is calibrated to obtain data at a sampling frequency of 25 Hz. Divide the 5 minute signal into 5

Each activity can obtain 480(60 × 8) signal segments[3].

3. Feature extraction

A total of 12 time-domain features were extracted: mean, standard deviation, root mean square, skewness, absolute mean, root mean square amplitude, kurtosis, maximum absolute value, margin index, waveform index, pulse index, peak index[4].

A total of 11 frequency domain indicators were extracted: mean value, variance, frequency root mean square, frequency center of gravity, root mean square frequency, combined feature 1, combined feature 2, combined feature 3, combined feature 4, combined feature 5[5]. The specific formula is shown in Table 1.

Table 1: Time - Frequency Domain Feature Extraction

Time domain		Frequency domain	
Method name	Formula	Method name	Formula
Average value	$TF_1 = \frac{1}{N}\sum_{i=1}^{N} x(i)$	Frequency mean	$FF_1 = \frac{1}{K}\sum_{i=1}^{K} y(i)$
Standard deviation	$TF_2 = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}[x(i)-\bar{x}]^2}$	Variance	$FF_2 = \frac{1}{K}\sum_{i=1}^{K} y^2(i)$
Root mean square	$TF_3 = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x^2(n)}$	Frequency root mean square	$FF_3 = \sqrt{\frac{1}{K-1}\sum_{i=1}^{K}[y(i)-FF_1]^2}$
Bias angle	$TF_4 = \frac{\sum_{i=1}^{N}[x(i)-TF_1]}{(\text{N-1})\sigma^3}$	The gravity center of frequency	$FF_4 = \frac{\sum_{i=1}^{K} f_i y(i)}{\sum_{i=1}^{K} y(i)}$
Mean absolute value	$TF_5 = \frac{1}{N}\sum_{i=1}^{N}\lvert x(i)\rvert$	Root mean square frequency	$FF_5 = \sqrt{\frac{\sum_{i=1}^{K} f_i^2 y(i)}{\sum_{i=1}^{K} y(i)}}$
Root amplitude	$TF_6 = \left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lvert x(i)\rvert}\right)^2$	Combination Feature 1	$FF_6 = \frac{\sum_{i=1}^{K}[y(i)-FF_1]^3}{K \times FF_2^{\frac{3}{2}}}$
Kurtosis	$TF_7 = \frac{\sum_{i=1}^{N}[x(i)-TF_1]^4}{(N-1)\sigma^4}$	Combination Feature 2	$FF_7 = \frac{\sum_{i=1}^{K}[y(i)-FF_1]^4}{K \times FF_2^{\frac{3}{2}}}$
Maximum absolute value	$TF_8 = \max(\lvert x\rvert)$	Combination Feature 3	$FF_8 = \sqrt{\frac{\sum_{i=1}^{K}[f_i - FF_4]^2 y(i)}{K-1}}$
Abundance index	$TF_9 = \frac{TF_8}{TF_6}$	Combination Feature 4	$FF_9 = \frac{\sum_{i=1}^{K}[f_i - FF_4]^3 y(i)}{(K-1)FF_8^3}$
Waveform index	$TF_{10} = \frac{TF_3}{TF_1}$	Combination Feature 5	$FF_{10} = \frac{\sum_{i=1}^{K}[f_i - FF_4]^4 y(i)}{(K-1)FF_8^4}$
Pulse indicator	$TF_{11} = \frac{TF_8}{TF_1}$	Combination Feature 6	$FF_{11} = \frac{FF_8}{FF_4}$
Peak indicator	$TF_{12} = \frac{TF_8}{TF_6}$		

ISSN 2616-5775 Vol. 6, Issue 6: 115-120, DOI: 10.25236/AJCIS.2023.060618

Since the subject 's body is equipped with a total of 45(5 × 9) sensors, and each sensor can extract 23 features, a total of 1035(45 × 23) features are extracted.

4. Dataset partitioning

The features extracted from the first 7 of the 8 subjects were used as the training set, and the action features extracted from the 8th subjects were used as the test set. Finally, 7980 rows of 1035 columns combined training set and 1140 rows of 1035 columns combined test set were sorted out.

5. Introduction of various machine learning methods

5.1 Logical regression

Logistic regression is an algorithm proposed to solve the binary classification problem. Both logistic regression and linear regression belong to generalized linear models. The core idea of the logistic regression algorithm is to use the known results to reversely derive the model that is most likely to cause the result.

Logistic regression has the advantages of fast speed, easy to understand, easy to update the model to absorb new data, but the application scenarios and data requirements are more stringent.

5.2 K-nearest neighbor algorithm (KNN)

The principle of KNN algorithm is to select the point with the smallest K in the feature space for complex analysis. If most of the K nearest points belong to a certain class, the sample is determined to be this class.

The advantages of KNN are: relatively simple, mature theory, insensitive to outliers. The disadvantage is: not suitable for large amount of data analysis, low performance.

5.3 Decision tree algorithm

The principle of the decision tree algorithm is to summarize a rule in the training set, so as to obtain a decision criterion that is less different from the training set.

The advantages of the decision tree algorithm are: fast classification speed, high model readability, easy to understand, and good generalization ability. The disadvantage is that it is easy to over-fit and ignore the correlation of the data set.

5.4 LightGBM algorithm

The LightGBM algorithm introduces two new technologies, GOSS and EFB, based on the GBDT algorithm.

LightGBM has the advantages of high speed, less memory and high precision. The disadvantage is that there may be a deep decision tree, resulting in overfitting.

6. Experiment

6.1 Data preprocessing

Before machine learning, this paper adopts normalization processing and PCA dimensionality reduction for feature data sets. Normalization processing can weaken the influence of outliers on data, and to a certain extent, it can accelerate the speed of gradient descent to find the optimal solution, and may improve the accuracy. PCA dimension reduction can avoid the dimension disaster caused by the high dimension of the data set, extract the main feature components of the data, and greatly improve the training speed of the model.

ISSN 2616-5775 Vol. 6, Issue 6: 115-120, DOI: 10.25236/AJCIS.2023.060618

6.2 Logistic regression algorithm

The original feature data set, the normalized feature data set, and the feature data set after PCA dimensionality reduction are used for training to obtain the effect of the model, as shown in Table 2.

Table 2: The training effect of logistic regression algorithm

Data set	Accuracy in the mean
Original feature data set	0.71
The normalized feature data set	0.89
The feature data set after PCA dimensionality reduction	0.89

The poor prediction effect of logistic regression algorithm is shown in Table 3.

Table 3: Logical regression algorithm predicts poor action

Data set	Action number	Precision	Recall	F1-score
Original feature data set	8	0.57	0.28	0.38
	9	0.61	1.00	0.75
The normalized feature data set	7	0.43	0.1	0.16
	8	0.58	0.68	0.63
The feature data set after PCA dimensionality reduction	7	0.38	0.18	0.25

6.3 KNN

The original feature data set, the normalized feature data set, and the feature data set after PCA dimensionality reduction are used for training to obtain the effect of the model, as shown in Table 4.

Table 4: The training effect of KNN algorithm

Data set	Accuracy in the mean
Original feature data set	0.67
The normalized feature data set	0.87
The feature data set after PCA dimensionality reduction	0.87

The poor prediction effect of KNN algorithm is shown in Table 5.

Table 5: KNN algorithm predicts actions with poor results

Data set	Action number	Precision	Recall	F1-score
Original feature data set	1	0	0	0
	2	0.19	0.87	0.31
	3	0	0	0
	4	0	0	0
	8	0	0	0
The normalized feature data set	2	0.53	0.92	0.67
	9	0	0	0
	11	0.48	0.92	0.63
The feature data set after PCA dimensionality reduction	2	0.54	0.92	0.68
	9	0	0	0
	11	0.46	0.85	0.6

6.4 Decision tree

The original feature data set, the normalized feature data set, and the feature data set after PCA dimensionality reduction are used for training to obtain the effect of the model, as shown in Table 6.

It can be known from the average accuracy that the effect of using the decision tree algorithm for human action classification is not ideal.

Table 6: Training effect of decision tree algorithm

Data set	Accuracy in the mean
Original feature data set	0.64
The normalized feature data set	0.54
The feature data set after PCA dimensionality reduction	0.62

ISSN 2616-5775 Vol. 6, Issue 6: 115-120, DOI: 10.25236/AJCIS.2023.060618

6.5 LightGBM algorithm

6.5.1 The original feature data and the normalized feature data set are used for training

Before training, the grid search is used to adjust the parameters to obtain the best parameters within a certain range. Under the premise of calling the best parameters, the LightGBM model is trained. The performance of the model is analyzed based on the results predicted using the test set[6, 7]. The specific data are shown in Table 7.

Table 7: Effects of LightGBM algorithm training

Data set	Accuracy in the mean
Original feature data set	0.77
The normalized feature data set	0.63

The poor prediction results obtained when using these two feature data are shown in Table 8.

Table 8: Training using the original data set and the normalized data set

Data set	Action number	Precision	Recall	F1-score
Original feature data set	2	0.22	0.03	0.06
	3	0	0	0
	7	0.25	0.3	0.27
The normalized feature data set	2	0	0	0
	7	0	0	0
	9	0	0	0
	11	0.13	0.07	0.09
	12	0.38	1	0.55
	18	0.25	1	0.4

6.5.2 Using PCA dimensionality reduction feature data for training

Table 9: Using PCA dimension reduction data set training

	Action label	Precision	Recall	F1-score	Accuracy
The feature data set after PCA dimensionality reduction	1	0.95	0.67	0.78	0.83
	2	0.68	0.28	0.4	
	3	0.95	1	0.98	
	4	1	1	1	
	5	0.94	1	0.97	
	6	0.94	0.73	0.82	
	7	0.27	0.28	0.27	
	8	0.49	0.73	0.59	
	9	1	0.35	0.52	
	10	0.66	0.98	0.79	
	11	0.94	0.98	0.96	
	12	0.87	1	0.93	
	13	0.91	0.97	0.94	
	14	0.86	1	0.92	
	15	0.85	1	0.92	
	16	1	0.77	0.87	
	17	1	1	1	
	18	0.87	1	0.93	
	19	0.94	1	0.97	

On the premise of calling the best parameters, the LightGBM model is trained using the feature data after PCA dimensionality reduction, which can greatly reduce the training time and improve the accuracy of the model to a certain extent. The specific training is shown in Table 9.

7. Analysis of experimental results

From the prediction results, the average accuracy of the logistic regression algorithm is only 0.71 when using the original feature data for training. The accuracy of 0.89 can be achieved by using the normalized data and the feature data after PCA dimensionality reduction, and the effect is ideal. However,

ISSN 2616-5775 Vol. 6, Issue 6: 115-120, DOI: 10.25236/AJCIS.2023.060618

the prediction effect of the normalized data on action 7 and action 8 is not very ideal, and the accuracy is 0.43 and 0.58 respectively. The prediction effect of the data after PCA dimensionality reduction on action 7 is not good, with an accuracy of only 0.38, a recall rate of 0.18, and an F1-score of 0.25.

For the KNN algorithm, the accuracy of using the original data is 0.67, and the accuracy is low. The accuracy of the model after normalization and PCA dimensionality reduction is 0.87, which is ideal. However, the prediction effect of the two at action 9 is not ideal, and the accuracy is 0. In addition, the prediction accuracy of action 2 and action 11 is low. Below 0.6, F1-score is below 0.7.

The effect predicted by the decision tree algorithm is not good. The average accuracy of the model after training using the original feature data set, the normalized feature data set, and the feature data set after PCA dimensionality reduction is 0.64, 0.54, and 0.62, respectively.

For the LightGBM algorithm, the accuracy of the prediction using the original feature data set is 0.77, the average accuracy of the model obtained using the normalized data is 0.63, and the accuracy of the feature data after dimension reduction using PCA is improved to 0.83. However, the data after dimension reduction using PCA has poor prediction effect on action 7 and action 8, with accuracy of 0.27 and 0.49 respectively, while the recall rate and F1-score of action 7 are also low, which are 0.28 and 0.27 respectively.

8. Conclusion

According to the above analysis, it can be known that the model trained by the feature data after PCA dimensionality reduction is generally better and the training speed is fast. For the application of human action classification, the logistic regression algorithm, KNN algorithm and LighttGBM algorithm are excellent in general, but there are great differences in the effect of some specific action classification. For example, the KNN algorithm with higher average accuracy cannot predict the action 9, while the LightGBM with an average accuracy of 0.83 has a very good prediction effect on the action, and the accuracy is 1.

The data in this paper are obtained by wearing inertial sensors. Although the required data can be accurately collected, this method of wearing instruments is not convenient in people's daily life. In order to collect data more conveniently, gyroscopes, acceleration sensors, and gravity sensors in smart phones can be used to collect data, and the data can be analyzed and processed accordingly to identify the actions of the owner.

References

[1] Huang Tao. Research on human motion recognition technology based on MEMS inertial sensors [D]. Harbin University of Commerce, 2021.
[2] Bao Yanyan. Research and application of machine learning in pose recognition [D]. Xi'an University of Architecture and Technology, 2018.
[3] Altun K, Barshan B, Tunçel O. Comparative study on classifying human activities with miniature inertial and magnetic sensors [J]. Pattern Recognition, 2010, 43(10): 3605-3620.
[4] Han Songshan. Research on human motion recognition based on multi-sensor joint [D]. University of Electronic Science and Technology of China, 2018.
[5] Zhuo S, Sherlock L, Dobbie G, et al. Real-time Smartphone Activity Classification Using Inertial Sensors—Recognition of Scrolling, Typing, and Watching Videos While Sitting or Walking[J]. Sensors, 2020, 20(3): 655.
[6] Jian Ding, Sun Yue. Multi-classification prediction of flight delays based on LightGBM [J]. Journal of Nanjing University of Aeronautics and Astronautics, 2021, 53(6): 847-854.
[7] Jin M , Zhang J , Huang T , et al. Research on Human Action Recognition Based on Global-Local Features of Video[C]// International Conference on Pattern Recognition and Machine Learning. IEEE, 2021.