Advances in Modern Agriculture (2022) Volume 3 Issue 2
doi: 10.54517/ama.v3i2.2051

ORIGINAL RESEARCH ARTICLE

Recognition of agricultural machinery operation trajectory based on BP_Adaboost

Yashuo Li[1,2,*], Weihao Fu[3], Xiaofeng Jia[1,2], Bo Zhao[1,2], Renzhong Su[4], Minghan Xu[1,2], Zaixi Pang[1,2], Yueru Zhang[1,2]

[1] *Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing 100083, China*

[2] *State Key Laboratory of Soil Plant Machinery System Technology, Beijing 100083, China*

[3] *China Machinery Construction Engineering Co., Ltd., Beijing 100083, China*

[4] *Center for Agriculture Mechanization Technology Extension of Hubei Province, Wuhan 430017, Hubei Province, China*

*** Corresponding author:** Yashuo Li, lys_caams@163.com

ABSTRACT

The movement track of agricultural machinery includes not only the field operation track but also the road driving track. Effectively distinguish the operation tracks of field operation and road driving during the driving of agricultural machinery, accurately divide effective operation plots, and accurately evaluate the operation efficiency of agricultural machinery so as to realize remote intelligent management of agricultural machinery. Typical characteristic data was extracted by analyzing attributes of agricultural machinery trajectory points, and a training model established by the method of BP_AdaBoost was used to recognize track points of agricultural machinery. After remarking error-prone track points at the junction of road and field, it was trained again. The correct rate of track recognition was 96.89%. It not only avoided the dependence of traditional clustering algorithms on thresholds and parameters but also effectively solved the problem of mistaking road driving trajectory into field operation trajectory.

Keywords: agricultural machinery operation track; BP_Adaboost; clustering algorithm

1. Introduction

With the improvement of the level of agricultural mechanization, the whole process of agriculture, such as deep loosening, plant protection, sowing, and harvesting, has been basically mechanized. However, due to the variety of agricultural machinery and the wide range of operations, it is difficult to supervise the quality of operation and evaluate the efficiency of operation, and other problems are becoming increasingly prominent. At present, the agricultural machinery information technology system has been able to realize the positioning, tracking, and remote monitoring of agricultural machinery operations. The location information of agricultural machinery can be used to extract driving tracks, and task allocation, operation evaluation, and subsidy settlement can be carried out according to the information on the Internet of Things[1–4]. Agricultural machinery driving trajectories are divided into road-driving trajectories and field-working trajectories. Due to non-standard operation of agricultural machinery or sensor problems, road transfer trajectories and field working

ARTICLE INFO
Received: 16 August 2022 | Accepted: 25 September 2022 | Available online: 28 October 2022

CITATION
Li Y, Fu W, Jia X, et al. Recognition of agricultural machinery operation trajectory based on BP_Adaboost. *Advances in Modern Agriculture* 2022; 3(2): 2051. doi: 10.54517/ama.v3i2.2051

trajectories during agricultural machinery driving cannot be effectively distinguished on the platform, which affects the evaluation of operation efficiency. At present, the trajectory point data processing mostly adopts a clustering algorithm, which can take advantage of the characteristics of dense field operation trajectory points and sparse road driving trajectory points and distinguish field operation and road driving trajectory by the DBSCAN clustering method, which can achieve a better identification effect[5,6]. Due to the large differences in the trajectories of different agricultural machinery and different terrain operations, it is difficult to determine the neighborhood radius and neighborhood points dependent on the clustering method, and the clustering method has a high time complexity. Some researchers use spatio-temporal information to perform density clustering on trajectories. Due to the temporal nature of trajectory points, taking time to cluster information is not helpful to improve accuracy[7,8]. Spatial index and grid density clustering can effectively improve the running speed, but the grid size and density threshold parameters are sensitive[9,10]. Although the clustering method has a good effect on the identification of agricultural machinery tracks in the field, it is easy to misidentify the road transfer tracks. It is difficult to improve the identification accuracy of agricultural machinery track points by simply relying on the clustering method.

In this study, the characteristics of the spatial running trajectory of agricultural machinery are combined, and the longitude and latitude, speed, and relationship between neighboring points of the agricultural machinery running trajectory are collected as features, and the BP_Adaboost method is used to train and recognize the trajectory points. Through the analysis of a large number of data points, it was found that the track points at the junction of road and field make it easy to misidentify the road driving track as the field operation track. Therefore, after BP_Adaboost makes the identification, the track points near the junction of field and road are marked as the road driving track and added to the training samples for repeated training until no new training samples are added. This method can effectively solve problems such as large differences in track in different plots and easy misidentification of track points at the junction of field roads.

2. Identification of driving track status of agricultural machinery

The running track status of agricultural machinery is mainly divided into field operation status, road running status, parking status (agricultural machinery failure, rest), etc. Generally, the Beidou positioning and receiving device installed on agricultural machinery automatically records the whole running track of agricultural machinery and obtains the longitude, latitude, position information, and speed information of agricultural machinery. In field operation, linear reciprocating operation is usually adopted, with slow movement speed and relatively dense trajectory points. When driving on the road, the moving speed is relatively fast, and the movement trajectory is a one-way linear distribution. When the farm machinery is in the state of parking, the characteristics of the farm machinery trajectory are cluster points, and the distribution range of cluster points is small and relatively scattered. By using these characteristics, the relationship between agricultural machinery track points is mined, and the running state of agricultural machinery can be effectively recognized.

2.1. Data preprocessing

The coordinates and velocities of agricultural machinery are the main characteristics of trajectory points. When the agricultural machinery stops, the on-board Beidou positioning and receiving device continues to receive information. Due to the positioning accuracy, there is a small deviation between the position of the trajectory point and the position of the agricultural machinery, which is scattered in a small area of the agricultural machinery position, which is called the agricultural machinery parking scatter. In special cases, due to excessive positioning deviation, the distance between two consecutive points will be very far, which is called trajectory point drift. Agricultural machinery parking scatter points and drift points are shown in **Figure**

1.

The effective trajectory points are continuous points in the normal driving process, and the agricultural machinery parking scatter points and the drift points in the driving process are invalid trajectory points, which will affect the calculation of average speed and trajectory spacing, and should be removed before data processing. In this study, in order to maintain the continuity and integrity of the trajectory, parking scatter points and drift points are still retained and classified as road driving points, but they are not included in the calculation during data processing.

Due to the error of track point collection, the track points will be concentrated in a small range when the farm machine stops, so the speed is zero or a small value. In order to prevent the temporary stop track points from being deleted by mistake during the field operation, the judgment condition was set as the speed of continuous multiple points less than a certain threshold.

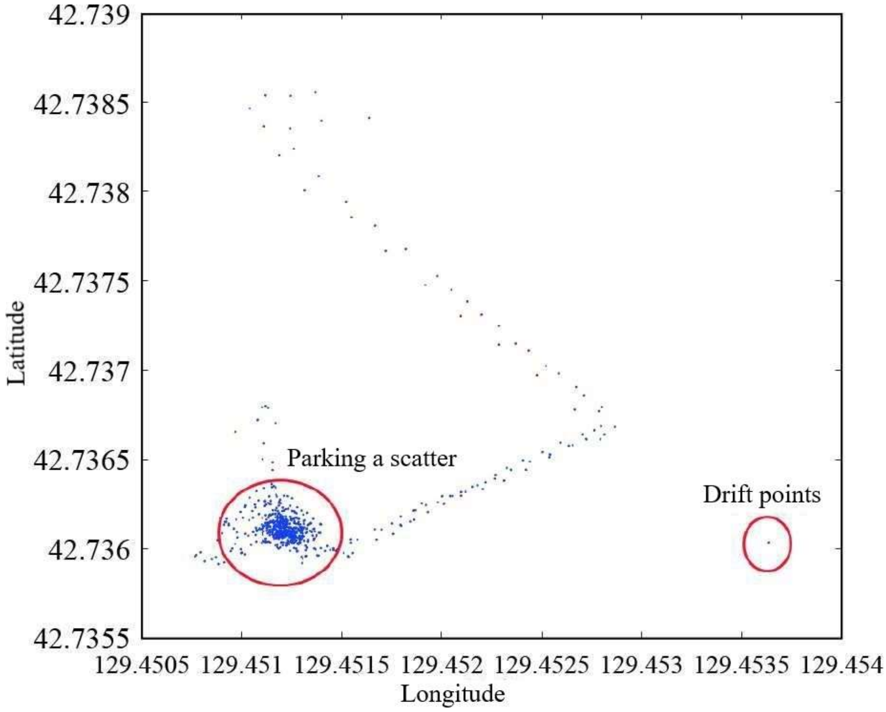


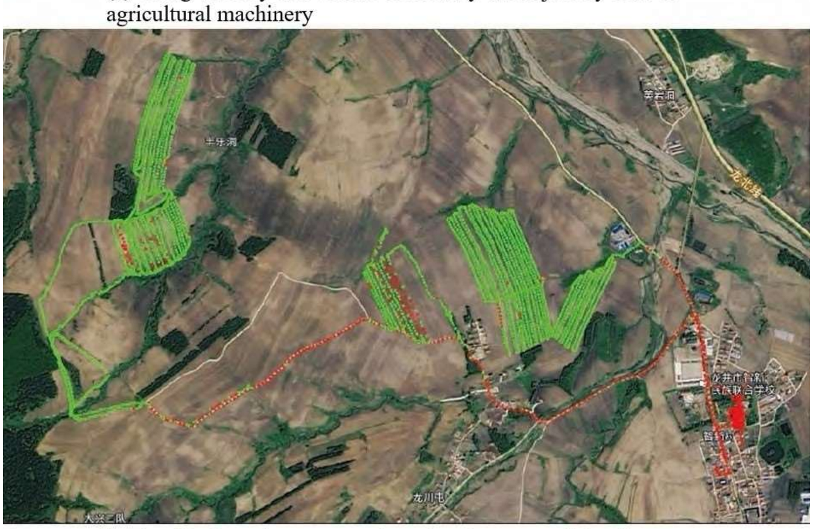
Figure 1. Scattered points and drift points of agricultural machinery parking.

Due to the failure of the Beidou positioning receiver, power failure, or positioning deviation, there will be a long distance between two consecutive points. When the distance between a point and the previous point is large, but the distance between the point and the next point is normal, it is considered that there is a breakpoint between the two points, and the previous point is regarded as the end point of the previous section of the trajectory, and this point is regarded as the starting point of the next section of the trajectory. If a certain point has a large distance from the two points before and after it, it is considered a drifting point and removed. The previous point is regarded as the end point of the previous section of the trajectory, and the next point is regarded as the starting point of the next section of the trajectory.

2.2. Feature selection

The track points of agricultural machinery contain information such as time, latitude and longitude, speed, etc. The track for agricultural machinery field work and the road running track are crossed. Clustering and classification methods are used to identify the state of track points according to the track point information, which can have certain effects. The trajectory recognition results, characterized by velocity and the mean distance between a trajectory point and 20 neighboring points, are shown in **Figure 2**.

(a) Using velocity as a feature to identify the trajectory state of agricultural machinery



(b) Identify the trajectory state by using the mean of the distance between the trajectory point and 20 neighboring points

Figure 2. Track effects recognized by different features.

In **Figure 2**, the points with dense reciprocating driving represent the field operation track, and the single driving point represents the road driving track. The selection of the two characteristics as the judgment basis has certain effects. Using the speed trajectory state judgment of agricultural machinery, the threshold for the average speed was set, but due to different velocity, field operation points were collected and road track point differences were bigger, and in the process of actual driving due to uneven agricultural machinery speed, especially in road deceleration, sudden acceleration lead to collect field data is not stable. Therefore, it is not good to distinguish the states of trajectory points by simply relying on velocity. When the distance relationship between track points and nearest neighbors is used, the distance between track points and nearest neighbors in the field is small and compact, while the distance between track points and nearest neighbors on the road is far. With the increase in the number of nearest neighbor points, the distance between the road trajectory points and the nearest neighbor points increases more. However, due to the possibility of deceleration and round-trips in the road driving process and the possibility of cross-ridge operation in the field, the track spacing becomes larger, which will lead to misidentified track points in the field as road and misidentified track points in the road as field.

The useful information about track points for agricultural machinery should be enriched as much as possible. The field operation of agricultural machinery is a round-trip operation, and the density of track points in field operation is greater than that of track points on the road. The available information includes velocity *V*, the number of neighboring points within a certain radius *M*, and the average distance *Aved* from neighboring points. The distribution effect of agricultural machinery trajectory points in the three-dimensional space composed of the above three features is shown in **Figure 3**.

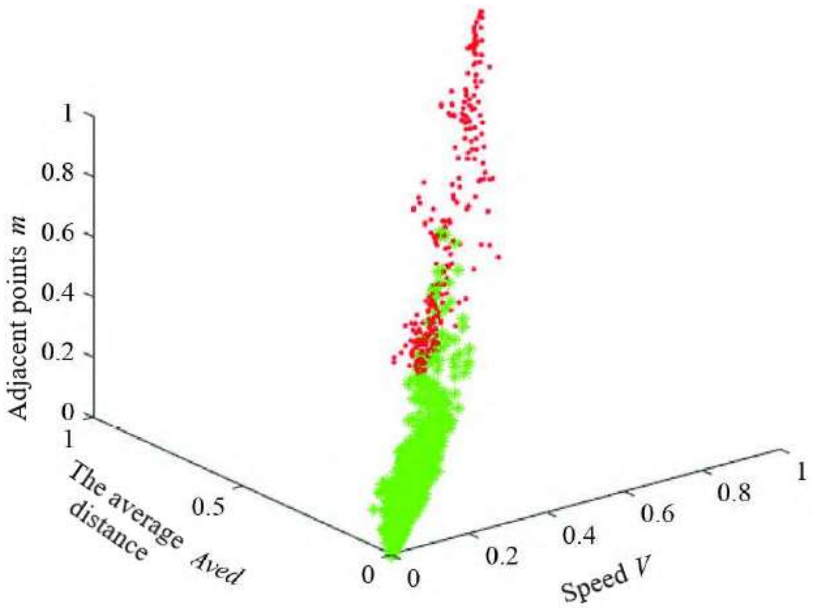


Figure 3. Effect drawing of track point distribution of agricultural machinery in three-dimensional space *V*, *m* and *Aved*.

In **Figure 3**, "*" represents the track point of field operation, and "•" represents road driving track point. The nearest neighbor points are selected to be 3 times the average distance of the daily trajectory points. As can be seen from **Figure 3**, although the two kinds of trajectory points are distinguishable in the 3D feature space, there are also overlapping parts that are difficult to distinguish. Therefore, by using the above features, the trajectory points have certain discriminability, indicating that the selected features are effective, but only relying on these several feature thresholds cannot completely distinguish the trajectory points. To sum up, the features selected in this paper include latitude and longitude, velocity, number of neighboring points, average distance of neighboring points, distance of N neighboring points, etc.

2.3. BP_Adaboost algorithm

The BP_Adaboost algorithm inputs the feature matrix into the BP classification algorithm based on adaboost for training. The idea of the Adaboost algorithm is to form a strong classifier for a given feature set by training a certain number of weak classifiers, and the weak classifier in BP_Adaboost is the BP neural network[11,12].

A BP neural network is composed of an input layer, a hidden layer, and an output layer, which is a mapping from input to output. By training the network with known patterns, a network that can reflect the exact mapping relationship between input and output can be obtained, which is characterized by signal forward propagation and error back propagation[13–15]. According to the existence theorem of mapping networks, a 3-layer feed-forward network can approximate any continuous function with any precision. In the modeling process, the initial learning rate is set to 0.0001, which can not only ensure that the minimum value of the loss function can be found quickly and smoothly but also avoid excessive repeated oscillation.

The input layer is the characteristics of the trajectory points, including the longitude and latitude of the trajectory points, the velocity, the number of neighboring trajectory points within the radius r, the average distance to the neighboring points, and the distance to the neighboring points. Since the units of each input variable are inconsistent and the value range differs greatly, normalization is carried out in advance according to the max-min principle[16] (Equation (1)).

$$L\left(\xi_{i,j}\right) = \frac{\xi_{i,j} - \xi_{i,min}}{\xi_{i,max} - \xi_{i,min}} \tag{1}$$

where $\xi_{i,j}$—the jth value of the input variable i; $\xi_{i,\min}$—The minimum value of the input variable i; $\xi_{i,\max}$—The maximum value of the input variable i.

The output layer contains two neurons, respectively corresponding to the trajectory point state. When the output of a neuron in the output layer is 1, it represents the farm machine operating in the field, and when the output is −1, it represents the farm machine driving on the road.

The number of neurons in the hidden layer can be determined as 8 according to Equation (2).

$$h = \sqrt{n_1 + n_2} + a \qquad 0 \leq a \leq 10 \tag{2}$$

where: h—number of neurons in hidden layer; n_1—Number of neurons in the input layer; n_2—Number of neurons; a—Variable coefficient.

In the input layer and hidden layer, the activation function of each neuron adopts tanh function

$$g(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3}$$

Specific steps: Firstly, an initial weight is set for each weak classifier, and then the weight is updated according to the error of each weak classifier on the feature set. The weight of the classifier with large error is increased, while the weight of the classifier with small error is reduced. After several cycles, the final weight obtained is combined with the corresponding weak classifier to obtain a strong classifier. BP_Adaboost algorithm consists of N BP neural networks, each BP neural network is a weak classifier, and N BP networks constitute a strong classifier. The processing flow is shown in **Figure 4**.

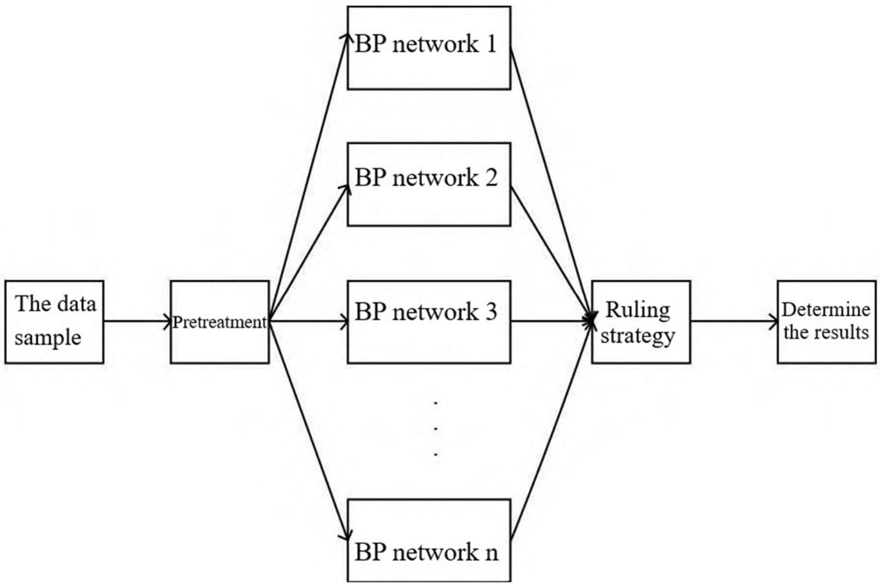


Figure 4. Flow chart of classification algorithm based on BP_adaboost model.

3. Test and results

The operation data of 50 farm machines from 15–26 October 2020 were used as the data set in the experiment. The Beidou positioning and receiving device is installed on the agricultural machinery, and the operation point data is collected every 5 s. The total number of trajectory points is 5196782, which is sent to the back-end server wirelessly. Each track point contains information such as the longitude and latitude of the farm machinery operation point and the speed of vehicles. In the experiment, longitude and latitude, speed, the number of nearest neighbors within 3 times the average distance, the distance of 20 points around the track point and the average distance were selected as the features. BP_Adaboost algorithm was used to train the model to identify the track point state of agricultural machinery. Because the junction of road driving track points and field work track points is easy to be identified incorrectly, and road driving track points are often identified as field work track points, this part is marked as road driving track points and put into training samples for re-training until no new training samples are added.

3.1. Test procedure

(1) Preprocessing the collected data points without calculating parking scatter points and abnormal drift points.

(2) The track points of field operation and road driving were marked as 1 and −1, respectively. 80% of the track points of each day were selected as training samples and all the track points of that day were selected as test samples.

(3) BP_Adaboost training dataset is used to identify the running state of test samples.

(4) Mark the track points with marked changes and the n track points before and after them as −1, and put these n +1 track points into the training sample.

(5) Repeat steps 3 and 4 until the number of training samples does not change.

3.2. Test results

According to the above test steps, the track point states of agricultural machinery are identified. The track point state recognition effect of one agricultural machinery in one day is shown in **Figure 5**.

In **Figure 5**, DBSCAN is used to cluster the track points of agricultural machinery. The number of points around the core point is 6, and the radius is 3 times the average distance. In the lower left corner of Figure 5A, dense track points in the area were misidentified as field work tracks due to the back and forth driving and steering sections in the road. Using single-BP neural network training, the error rate is high, and many road driving tracks are identified as field work tracks. In this paper, 20 BP neural networks were selected and trained by the BP_Adaboost method. The track points with changed marks and nearby points were marked as road track points and re-trained with training samples, which could solve the problem of wrong recognition of the intersection of some road tracks and field roads, but there were still some wrong recognitions of some areas. The three methods perform state recognition for all trajectory points, and the accuracy is shown in **Table 1**.

(a) DBSCAN method

(b) BP neural network

(c) BP_Adaboost

Figure 5. Identification effect of each method.

The number of nearest neighbor points in the feature of trajectory points is denoted as f_{t_n}, and the number of change points marked by trajectory points and the number of points selected before and after are denoted as f_{g_n}. The selection range of f_{t_n} is 3–18, and the selection range of f_{g_n} is 7–25, respectively. The results are shown in **Figure 6**.

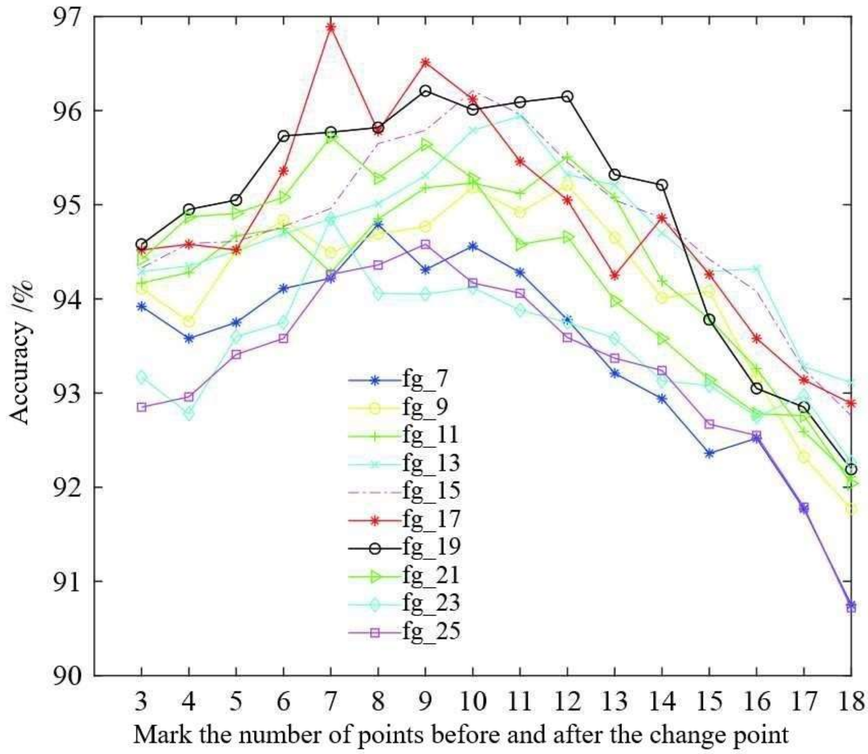


Figure 6. Identification effect drawing of different values of f_{t_n} and f_{g_n}.

Table 1. Recognition accuracy of each method.

Methods	**DBSCAN**	**BP**	**BP_Adaboost**
Accuracy /%	93.35	92.62	96.89

The experimental results show that, on the general trend, the classification accuracy will improve with the increase of the number of nearest neighbors, and the accuracy starts to decline when it exceeds 7. When tested by BP_Adaboost, with the increase in the number of before and after the retrained marker change track points, the classification accuracy will also improve, and when it exceeds 17, the accuracy will decrease. The main reasons are as follows:

(1) The track points of agricultural machinery field work are relatively clustered, while the track points of road driving are generally single and discrete. When the number of adjacent points increases, the sample information is richer and the training effect is better. When the number of adjacent points increases too much, the number of points near the road driving point also increases, and the difference between them and the track points of field work decreases, which leads to a decline in classification accuracy.

(2) When BP-Adaboost identified the test samples, most of the errors occurred in the places where the marks of driving track points changed, and the road driving points were identified as the track points of field operations. The adjacent points before and after the trajectory points whose marks have changed are marked as road driving points, which are put into the training samples for re-training, and the test samples are classified again, which can not only enrich the training samples but also focus on the error-prone trajectory points. Therefore, with repeated training and testing, the classification accuracy will improve. When the number of former nearest neighbors increases, the training speed is faster, the number of samples is larger, and the accuracy is improved accordingly. If it is too large, it is easy to mislabel the field operation track points as road driving track points, and the additional wrong training samples will affect the training model and then affect

the recognition accuracy. In some farm machinery operations, the selection is too large, and even the recognition accuracy will fall to less than 70%.

4. Conclusion

The state recognition of agricultural machinery running track can effectively evaluate the working efficiency of agricultural machinery. In this paper, the longitude and latitude, velocity, and neighboring point relationships of agricultural machinery track points are selected as features, and the BP_Adaboost method is used to establish a training model to identify agricultural machinery track points. By analyzing the results, the track points near the junction of road and field that are easy to misidentify are marked as road driving tracks and added to the training samples for repeated training until the training samples are no longer increased. The identification accuracy of this method is 96.89%. After effectively distinguishing field operation track and road driving track, field operation track can be better divided by plot, and then the operation efficiency of each plot can be evaluated. Especially for some farm machinery operations subsidized according to plots, the division of plots and effective assessment of operation quality can make operation supervision more efficient and accurate. Although the accuracy is improved compared with the DBSCAN clustering algorithm and the BP neural network method alone, the number of neighboring points in features and the selection of points before and after the change of markers will affect the results, and the sample labeling is time-consuming. Therefore, finding a stable classification method and reducing the workload of labeling is the focus of future research.

Conflict of interest

The authors declare no conflict of interest.

References

1. Zhang F, Liu R, Ni Y, Wang Y. Dynamic positioning accuracy test and analysis of Beidou satellite navigation system (Chinese). *GNSS World of China* 2018; 43(1): 43–48.
2. Zheng Y, Zhao Y, Liu W, Liu S. Forest microclimate monitoring system based on Beidou satellite (Chinese). *Transactions of the Chinese Society for Agricultural Machinery* 2018; 49(2): 217–224.
3. Li D. Internet of things and smart agriculture (Chinese). *Agricultural Engineering* 2012; 2(1): 1–7.
4. Chen Y. Discussion on the relationship between agricultural Internet of things and intelligent control of agricultural complex large-scale systems (Chinese). *Agricultural Network Information* 2012; 2: 8–12.
5. Liu H, Qiao Y, Zhao G, et al. Agricultural machinery abnormal trajectory recognition. *International Journal of Machine Learning and Computing* 2021; 11(4): 291–297. doi: 10.18178/ijmlc.2021.11.4.1050
6. Marmat T, Xie J. Research on clustering of agricultural machinery operation trajectory based on DBSCAN algorithm (Chinese). *Journal of Agricultural Mechanization Research* 2017; 39(4): 7–11.
7. Wu D, Du Y, Yi J, et al. Density-based spatiotemporal clustering analysis of trajectories (Chinese). *Journal of Geo-information Science* 2015; 17(10): 1162–1171.
8. Shaw S, Yu H, Bombom LS. A space-time gis approach to exploring large individual-based spatiotemporal datasets. *Transactions in GIS* 2008; 12(4): 425–441. doi: 10.1111/j.1467-9671.2008.01114.x
9. Wang P, Meng Z, Yin Y, et al. Automatic identification algorithm and experiment of field operation status based on spatial trajectory of agricultural machinery (Chinese). *Transactions of the Chinese Society of Agricultural Engineering* 2015; 31(3): 56–61.
10. Jokinen J, Räty T, Lintonen T. Clustering structure analysis in time-series data with density-based clusterability measure. *IEEE/CAA Journal of Automatica Sinica* 2019; 6(6): 1332–1343. doi: 10.1109/jas.2019.1911744
11. Li N. Cheng X, Zhang S, et al. Recognizing human actions by BP-AdaBoost algorithm under a hierarchical recognition framework. In: Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 26–31 May 2013; Vancouver, BC, Canada. pp. 3407–3411. doi: 10.1109/icassp.2013.6638290
12. Li B, Zhang X, Fang H. An improved BP-Adaboost algorithm and its application in radar multi-target classification (Chinese). *Journal of Nanjing University (Natural Science)* 2017; 53(5): 984–989.
13. Jiang Y, Zhang F. Study on BP neural network optimization by improved decay parameter genetic algorithm. *Journal of Physics: Conference Series* 2020; 1621(1): 012054. doi: 10.1088/1742-6596/1621/1/012054

14. Dai Z, Yao L, Qin J. Research on fault prediction of radar electronic components based on analytic hierarchy process and BP neural network. In: Proceedings of the 2019 12th International Conference on Intelligent Computation Technology and Automation; 26–27 October 2019; Xiangtan, China. pp. 91–95. doi: 10.1109/icicta 49267.2019.00026
15. Dong M, Jiang T, Yan W, et al. Research on the application of BP neural network and decision tree technology in data mining (Chinese). *Economic Research Guide* 2018; 20: 186–190.
16. Bao H. *Energy Management Strategy of Roadside Supercapacitor Energy Storage System Based on Energy Prediction of BP Neural Network* (Chinese) [Master's thesis]. Beijing Jiaotong University; 2020.