Automation and Machine Learning (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/autml.2025.060107
ISSN 2516-5003 Vol. 6 Num. 1

Micro motion detection system for traffic target classification

Zhangxiaoyu Wu*, Jiacong Guo, Hang Zhou

School of Mechanical and Electrical Engineering, Wuhan University of Technology, Wuhan, 430070, China
**Corresponding author*

Keywords: Intelligent transportation, micro motion detection, time-frequency characteristics, multi-layer perceptron

Abstract: The main way of traffic monitoring is through cameras, but it is easily affected by external lighting. The ground is subject to various natural and human movements, which generate small vibration signals. Using micro motion signals to perceive the environment is a new detection technology. This research design consists of micro motion acquisition hardware, data processing and time-frequency domain feature extraction algorithms, and machine learning based object classification algorithms to achieve scene perception of people and vehicles in smart transportation, and can also be used for security detection. After collecting field data, the 2935 effective slices extracted by the algorithm were trained and tested. The average accuracies of the MLP, SVM, GBT, and RF models were 95.808%, 91.039%, 95.570%, and 95.468%, respectively. The MLP (Multi-Layer Perceptron) with shallow neural network structure is the optimal model, with the highest average recognition rate and the smallest standard deviation. The experimental results show that this system is not affected by environmental factors such as weather, light intensity, and electromagnetic fields. It is compact and easy to deploy, with strong concealment, small data volume, and high reliability.

1. Introduction

The 14th Five Year Plan for Technological Innovation in the Transportation Sector clearly proposes to vigorously develop smart transportation and promote the integration of new generation information technologies such as cloud computing and big data with transportation. As a key technical tool, the traffic detection system can monitor road traffic conditions in real time by collecting and analyzing traffic data, providing important information for urban traffic management departments and vehicle drivers, thus better responding to traffic congestion and reducing the incidence of traffic accidents, thereby improving the overall operational efficiency of the city and the travel experience of residents. Detecting traffic targets through cameras is the mainstream direction. Traditional methods include differential image method, edge detection method, optical flow method, etc. [1], while emerging methods rely on deep learning, but are still heavily influenced by lighting (night, strong light, dust, rain, snow, smoke). In recent years, Tao Jiadong [2] has designed a point to gravity weighted vehicle ranging method around the centroid based on the R-CNN algorithm, which has

improved detection performance by 2%, but the average error is still above 5%. Peng Hongxing et al. [3] proposed an improved YOLOv5 algorithm, which achieved an average detection accuracy of 95.7% for vehicles and personnel in highway tunnels. Feng Xiaomei [4] added a false detection module based on deep learning and improved the loss function of the detection algorithm. In three types of traffic target tests, the average accuracy reached 97.34%, 87.24%, and 80.23%. It can be seen that although visual algorithms are constantly developing, the difficulty of being easily affected by lighting still exists.

Traffic target detection based on the principle of small vibrations is a novel method, which has been extensively studied by Ahmad Bahaa Ahmad et al. They trained 5580 unprocessed vibration data using DNN, CNN, and RNN neural networks, and concluded that the CNN model is the most reliable [5]; Multiple artificial intelligence methods were used to evaluate 4185 samples of buses, cars, and motorcycles, with accuracy ranging from 62% to 97% [6]. Luo Zhiyuan et al. [7] constructed 1dCNN and 2dCNN axis information recognition network models, achieving recognition accuracies of 98.6% and 97.3% on 1229 test sets of two to six axis vehicles, respectively. It can be seen that traffic target detection and classification based on micro motion principle is feasible. This study is based on micro motion detection, designing data acquisition hardware, and creating datasets including personnel, two wheelers, and four wheelers. Through feature extraction and classification algorithms, a recognition rate of over 95% is achieved, providing a systematic reference scheme for traffic target classification.

2. Hardware solution

2.1 Overall hardware solution

The hardware part consists of three parts: micro sensors, hardware circuits, and plastic casing. The three tail cone plugs of the microsensor make contact with the ground and use ground vibration signals to obtain electrical signals along the X, Y, and Z axes (the X axis is a straight line parallel to the handheld handle segment on the ground plane, the Y axis is a straight line perpendicular to the handheld handle segment on the ground plane, and the Z axis is a straight line perpendicular to the ground plane); The hardware circuit performs preliminary processing on the signals collected by the sensors and stores the data; The device casing stably connects and protects the sensors and circuits, reducing damage caused by collisions. There is a handle directly above the device casing, which is easy to carry and deploy quickly. The overall appearance of the hardware is shown in Figure.1

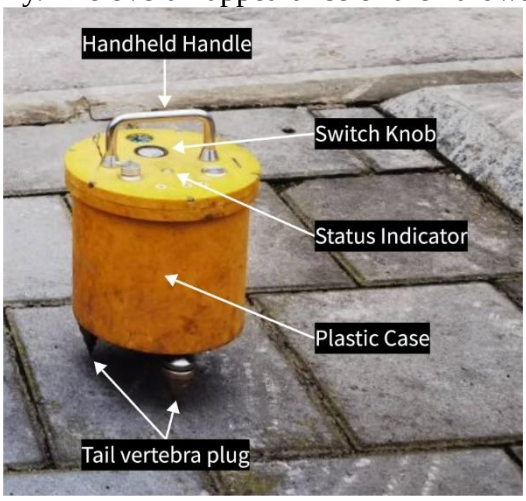


Figure.1 Overall appearance of hardware

2.2 Selection of sensor

The performance of the sensor is very important for data acquisition. In this study, PS-4.5C1 low-frequency seismometer is used to collect the micro-motion signal. With high sensitivity and rotating coil structure, the sensor can eliminate lateral impact force and selectively detect longitudinal wave and shear wave, meeting the application requirements of geophysical prospecting, engineering vibration measurement and other vibration measurement fields. The specific parameters are shown in Table 1.

Table 1 PS-4.5C1 Specific parameters of low-frequency seismometer

PS-4.5C1 Features	Parameters (20℃)
Frequency (Hz)	4.5±0.5
Sensitivity (V/m/s)	82±7.5%
Open Circuit Damping	0.58±20%
Coil Resistance (Ω)	3400±10%
Inertial Mass (g)	11.3
False Frequency (Hz)	≥90
Maximum Limit (mm)	4
Distortion (%)	≤0.3
Size (mm)	25.4×36
Temperature Range (℃)	-40 ~ +70
Weight (g)	80

2.3 Hardware circuit

The main function of the hardware circuit is to preliminarily process the analog voltage signal input by the sensor, reduce the natural frequency of the sensor to below 0.45Hz through the low-frequency expansion circuit and amplify the effective signal. The analog voltage is converted into digital signal through the AD conversion circuit, and the digital signal is sent to the upper computer through the serial port by the single-chip computer and stored in the SD card for saving. The LED indicator shows the working state of the hardware. The hardware circuit design structure is shown in Figure.2

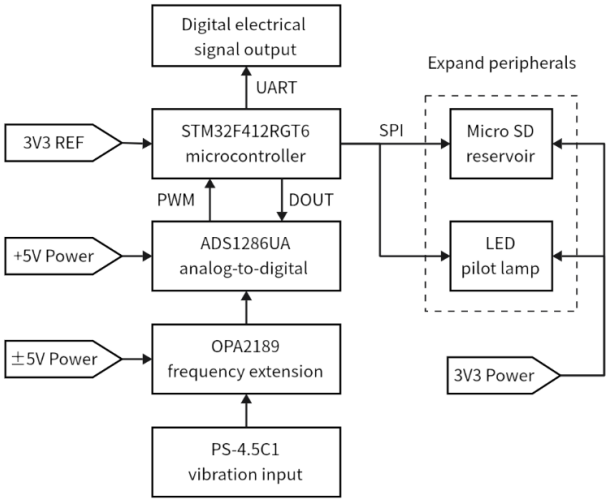


Figure.2 Hardware circuit design structure

3. Data acquisition

3.1 Collection scheme

The vibration frequency of personnel and vehicles is almost below 100Hz [8]. To avoid undersampling, Nyquist's sampling law shall be observed, i.e. the sampling frequency is more than twice the system frequency. However, in practice, the sampling frequency is higher. The team sampled at a frequency of 500Hz and collected 6 signals including cars, bicycles, electric vehicles, pedestrians, runners and the environment. The environmental vibration was collected for 60 minutes and the other 5 targets were collected for 500 times respectively. The data collected by the hardware is subject to data processing and feature extraction before target classification.

The quantity and quality of data sets are important factors for the accuracy rate of target classification. Therefore, 4 sets of hardware systems are carried during the actual acquisition and tested in turn. In this way, it can not only judge whether the hardware equipment has fault by checking the measured data, but also collect more data samples under the same test times to reduce outdoor workload.

3.2 Field collection

During field acquisition, different targets will pass through acquisition points at the same time, and vector superposition of vibration waves will occur [9]. In order to obtain reliable sample data, variables shall be controlled to obtain pure vibration data of single target as far as possible before exploring vibration of multiple targets. Firstly, find an area with minimum external interference and collect sufficient vibration data of personnel and two-wheel vehicle. Find a relatively ideal road (moderate traffic flow, no long-term blocking of vehicles and no passing of vehicles for a long time) and collect vibration data of cars. Field data collection is shown in Figures.3 (a), (b), (c) and (d).

(a) Walking (b) Running (c) Electric vehicle (d) Car

Figure.3 Field data acquisition diagram

4. Data processing

4.1 Data pre-processing

Through field collection, 630.8 MB of original vibration data is obtained. The data is stored in a. txt file in numerical form. The numerical value is the output voltage of the sensor, in mV. Positive and negative represent the vibration direction, and are retained to 6 decimal places. To make the value more intuitive, first convert the voltage into an amplitude in mm. For speed type moving coil sensors:

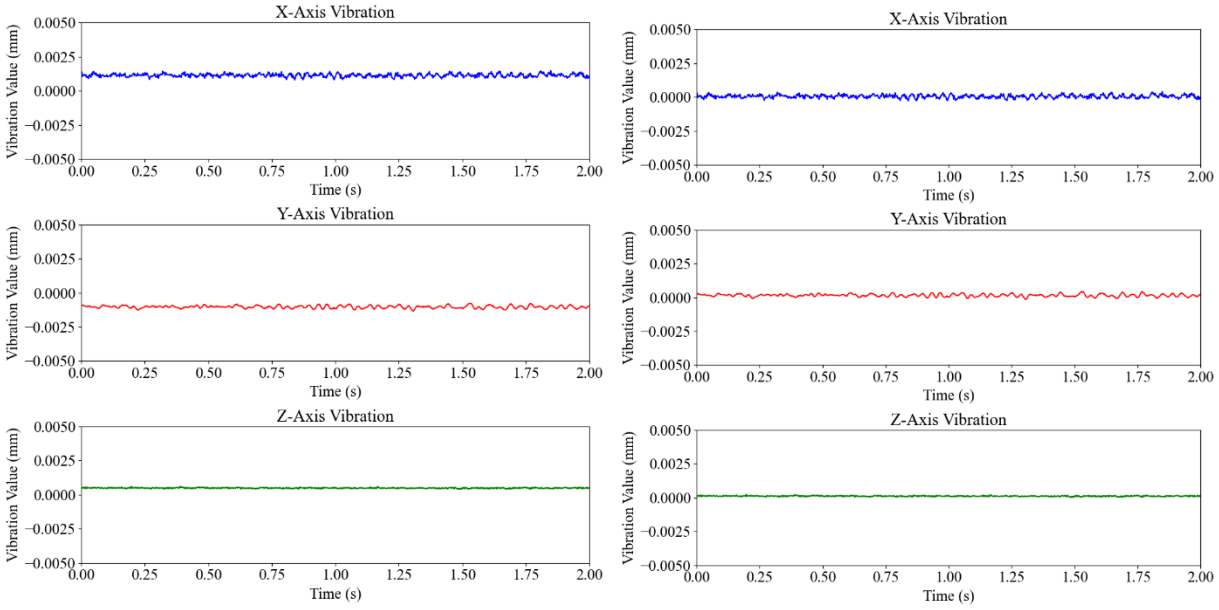
$$v = \frac{V}{S} \ , \ A = \frac{v}{2\pi F} \tag{1}$$

In formula (1), v is vibration velocity of sensor, V is output voltage of sensor, S is sensor sensitivity, A is vibration amplitude of sensor and F is vibration frequency of sensor. The sensor sensitivity

(V/m/s) used in this study was 82 ±7.5% and the vibration frequency was 4.5 ±0.5 Hz, which were converted.

Ground vibration waves are often accompanied by more low-frequency noise [10], and some components in the hardware circuit, especially analog amplifiers, will cause zero drift. For this purpose, filtering of all vibration data is necessary. In this study, a second-order high-pass Butterworth filter [11] is used, which attenuates below the cut-off frequency and can filter out low frequency noise and DC bias; No extra phase shift is introduced to the passing signal, and the signal phase will not be changed.

The time-vibration diagram of environmental test before and after filtering is shown in Fig.4. The horizontal axis represents time, the vertical axis represents amplitude, and blue, red and green curves represent vibration in X, Y and Z directions respectively. It can be seen from the figure that the unfiltered vibration data has a clear zero drift with an offset of 0.001620mm on the X axis. The reference line of vibration wave returns to zero scale after filtering.



(a) Unfiltered (b) Filtered

Figure.4 Time-vibration diagram for environmental testing

The vibration data collected is a long-term continuous signal. In order to analyze the specific vibration target in time and frequency domain, the vibration data must be sliced. To ensure consistent slice length and extract key features, an adaptive slice algorithm was designed in this study. Firstly, detect the maximum value. In a certain time range, the vibration value has a maximum value, which is greater than all other vibration values in the time range. When the maximum value is found, the initial slice of the vibration data in a certain range before and after is carried out with it as the center. Secondly, set secondary judgment conditions to eliminate environmental vibration interference. The ambient vibration wave is relatively gentle, while the vibration wave with specific target has large fluctuation. Therefore, a threshold is set. When the standard deviation of a slice is greater than the threshold, it is considered that the slice has obvious fluctuation characteristics, and it is an effective slice. Conversely, it is considered as an invalid slice and is removed. Finally, 2935 valid slices were obtained.

4.2 Time-domain feature analysis

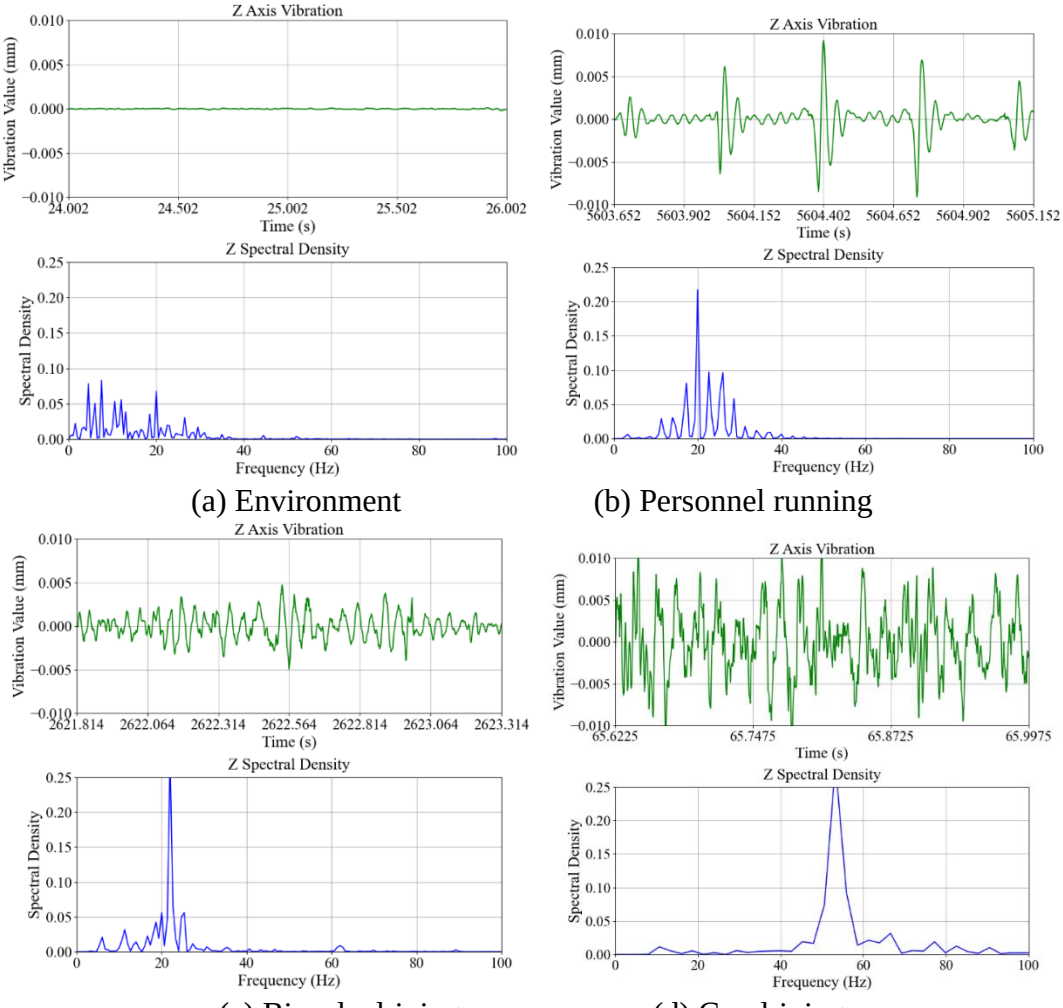
Time-domain feature analysis studies the change rule of signal form with time, extracts typical

feature quantities such as amplitude, period, local rising time and falling time as important basis for signal judgment and identification.

Frequency-domain energy of signal [12] refers to the energy of signal calculated by all or part of frequency points of signal in the whole or certain frequency band. Convert the signal to the frequency domain by FFT transformation, and calculate the frequency domain energy of the signal, the calculation formula is as follows(2):

$$E_f = \sum_{i=1}^{N} X(i)^2, \quad X(i) = \sum_{n=1}^{N} x_n e^{-j\frac{2\pi}{N}ni}, i = 1,2,\ldots,N-1 \tag{2}$$

Since there are many interferences between the vibration data of X axis and Y axis on the surface [13], only the vibration data of Z axis vertical to the surface plane is used in the subsequent calculation. Calculate time-frequency characteristics of each effective slice, including 7 time-domain characteristics: maximum value, minimum value, average value, standard deviation, peak factor, pulse factor, energy and 6 time-domain characteristics: main frequency, spectrum peak value, number of peak values, spectrum bandwidth, spectrum density and spectrum peak width. The 13 features of each active slice are stored in an Excel table for subsequent data analysis. The graphical results of the four typical features are shown in Figure.5, (a) for the environment, (b) for running, (c) for bicycles, and (d) for time-frequency features of cars, with the green curve above representing the time-vibration plot and the blue curve below representing the spectrum density plot.



(a) Environment (b) Personnel running

(c) Bicycle driving (d) Car driving

Fig.5 Extraction results of four typical features

The analysis of the image shows that the environmental vibration waveform is very gentle, close to the zero scale line, and the spectrum is chaotic; The running waveform of the personnel has a large vibration at the fall point of the foot, and a small aftershock during the stagnating period of the foot, with the main frequency of about 19Hz; As the bicycle passes through the collection point, the waveform is continuous and increases first and then decreases, and the main frequency is between 18-25Hz; The speed of the car is fast, the waveform amplitude is large and continuous, and the main frequency is 40-80Hz. The intuitive and easily distinguishable image results show that the data acquisition and processing have achieved good results.

5. Target classification

Since the time-frequency domain characteristics have been extracted, which is equivalent to the convolution operation of deep learning [14], this study uses the machine learning model for judgment verification, and uses four models: multilayer perceptron (MLP), support vector machine (SVM), gradient lifting tree (GBT) and random forest (RF). Of 2935 valid slices are used for training and 20% for testing, and the accuracy rate is based on whether the predicted result of the program is the actual result of the test set. Since the training set and test set are completely random, the accuracy will fluctuate slightly even if the same program is run. Therefore, each model was run ten times with appropriate parameters and the average accuracy and standard deviation were calculated to obtain more accurate experimental results. The ten test results of the four models are shown in Table 2, and the accuracy comparison diagram of the four models is shown in Figure.6:

Table 2 Ten test results of four models

Name of algorithm	Minimum accuracy rate	Maximum accuracy rate	Mean ten accuracy	Standard deviation
MLP	95.23%	96.59%	95.808%	0.00528
SVM	89.61%	93.53%	91.039%	0.01231
GBT	94.55%	97.27%	95.570%	0.00928
RF	94.21%	96.93%	95.468%	0.00840

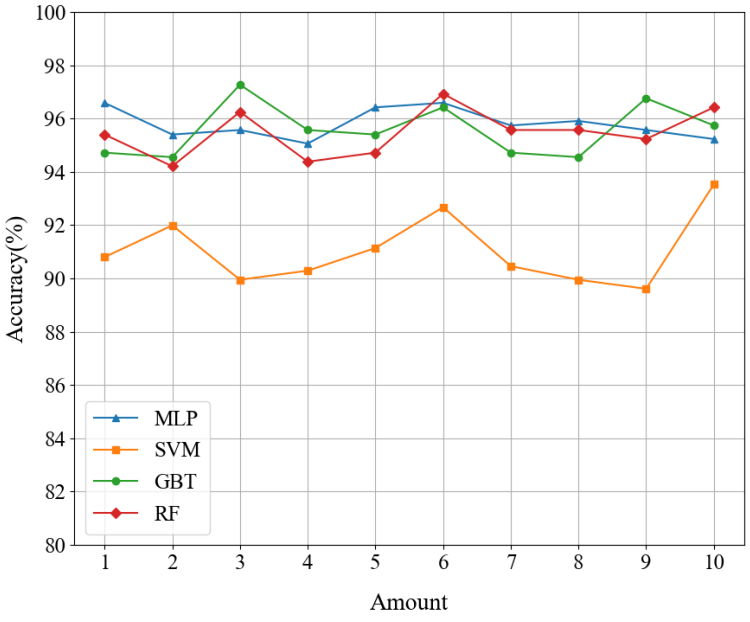


Figure.6 Accuracy comparison of four models

The ten-time average accuracy of MLP, SVM, GBT and RF models is 95.95.95.08%, 91.039%, 95.570% and 95.468% respectively, which indicates that the model achieves high recognition rate. Among them, GBT achieves the highest accuracy rate of 97.27% in a single shot, and SVM achieves the lowest accuracy rate of 89.61% in a single shot. The average accuracy of SVM is slightly lower than that of the other three models, because SVM has high requirements for data feature matching, its performance is greatly affected by parameter selection, and it is sensitive to data noise and outliers. GBT and RF are integrated learning algorithms, which improve the accuracy and generalization ability of the model by training multiple decision trees, with strong robustness. MLP is a shallow neural network model composed of multiple neuron layers with strong nonlinear fitting ability. In this study, it consisted of 2 hidden layers, each consisting of 100 neurons. The average accuracy of the model is the highest, and the standard deviation is the smallest, which is the optimal solution.

Taking the optimal solution multi-layer sensor as an example, the target category recognition rate in the test set is calculated, as shown in Fig.7. From the histogram, it can be seen that the identification rate of personnel walking, personnel running and car driving is 100%, because the vibration time-frequency domain characteristics of these three types of targets are significantly different from other types. The identification rate of environmental vibration is 99.3%, which may be caused by vibration disturbance caused by vehicle driving at a distance. However, the recognition rate of bicycle driving and electric vehicle driving is relatively low, 83.52% and 93.86% respectively, because the motion mode and speed of the two are relatively close, and the time-frequency domain characteristics have a certain intersection. Specifically, the vibration amplitude and wave peak number of the two are close to each other, the main frequency of bicycle driving is 13-28Hz, and the main frequency of electric vehicle driving is 18-33Hz.

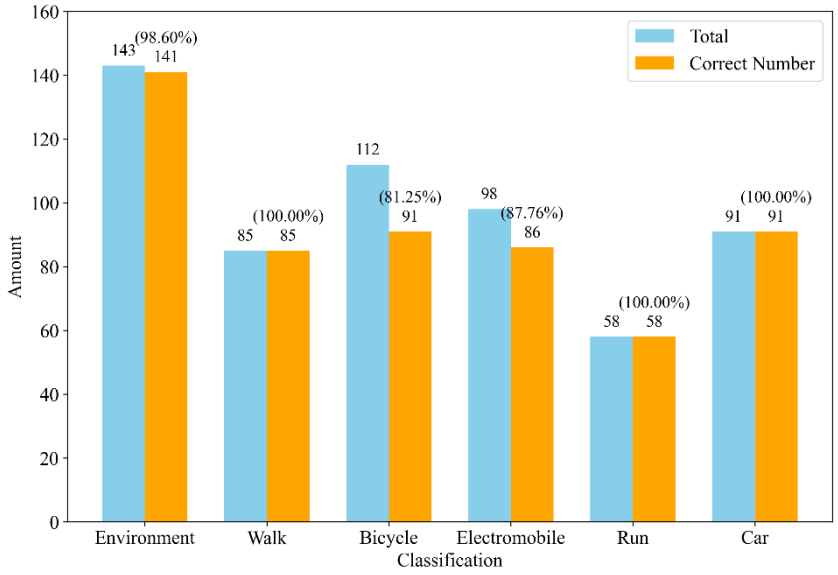


Figure.7 Histogram of target category recognition rate of MLP model

6. Conclusion

In the classification task oriented to traffic target, although the mainstream vision algorithm is developing continuously, it is still affected by light and other factors in the complex reality. In this paper, a set of vehicle detection system is designed based on the principle of micro-vibration. 2935 traffic target data sets including environment, personnel walking, personnel running, bicycles, electric vehicles and cars are collected in the field. After 13 time-frequency characteristics are extracted, four machine learning models are tested, and the average accuracy of the model for ten times is more than

90%. Among them, the Multilayer Perceptron (MLP) with shallow neural network structure is the optimal model, with the average accuracy rate of ten times reaching 95.08% and high recognition. At the same time, when the training set and test set are completely random, the standard deviation of the model for ten runs is only 0.00528, which shows strong stability. The system can be directly buried on both sides of the road and connected to the network cable to collect the information of people and vehicles on the whole road. It is almost unaffected by the weather, light and electromagnetism and has high reliability. It can also be used for security detection and is not easy to be found or damaged. In addition, the vibration data acquired by the system occupies a small amount of memory, and the data volume of a single device in one hour is about 48MB, which has certain advantages over the huge data volume of the camera.

Acknowledgements

The authors gratefully thanks for funding the National Innovation Training Program for University Students 2023 (S202310497157).

References

[1] Xiaoyi W, Yunhua L, Zhongjun Y U, et al. Multi-channel radar target micro-motion feature classification method based on EfficientNet[J].Systems Engineering & Electronics, 2024, 46(9).DOI:10.12305/j.issn.1001-506X.2024.09.11.

[2] Richter Y, Balal N. High-Resolution Millimeter-Wave Radar for Real-Time Detection and Characterization of High-Speed Objects with Rapid Acceleration Capabilities[J].Electronics, 2024, 13(10):13.DOI:10.3390/electronics13101961.

[3] Ahmad B I, Harman S, Godsill S.A Bayesian track management scheme for improved multi‐target tracking and classification in drone surveillance radar[J].IET Radar, Sonar & Navigation (Wiley-Blackwell), 2024, 18(1). DOI:10.1049/rsn2.12458.

[4] Wang X, Zhang M, Li B.Fusion Network Based on Motion Learning and Image Feature Representation for Micro-Expression Recognition[C]//Chinese Conference on Pattern Recognition and Computer Vision, (PRCV).Springer, Singapore, 2025.DOI:10.1007/978-981-97-8795-1_37.

[5] Voronin V, Zhdanova M, Semenishchev E,et al.Action recognition algorithm from multimodal images for human-robot collaboration systems[J].Proceedings of SPIE, 2023, 12621(000):9.DOI:10.1117/12.2678112.

[6] Voronin V, Zhdanova M, Tokareva O,et al.A multimodal visual guided robot collaborative system based on the classification of multi-class human motion[J].Proceedings of SPIE, 2023, 12766(000):7.DOI:10.1117/12.2691154.

[7] Huan S, Wu L, Zhang Z Y C.Radar Human Activity Recognition with an Attention-Based Deep Learning Network[J].sensors, 2023, 23(6). DOI:10.3390/s23063185.

[8] Hassan S, Wang X, Ishtiaq S,et al.Human Activity Classification Based on Dual Micro-Motion Signatures Using Interferometric Radar[J].Remote Sensing, 2023, 15(7).DOI:10.3390/rs15071752.

[9] Ryu B, Kwak D.Site effects and micro-zonation 1 Development of Nonlinear Site Amplification Model Based on Average Shear-wave Velocity of Soil Layer and Bedrock Depth[J].Japanese Geotechnical Society Special Publication, 2024, 10(28):5.DOI:10.3208/jgssp.v10.OS-17-05.

[10] Ashkenazi I, Benady A, Zaken S B,et al.Radiological Comparison of Canal Fill between Collared and Non-Collared Femoral Stems: A Two-Year Follow-Up after Total Hip Arthroplasty[J].Journal of Imaging, 2024, 10(5). DOI:10. 3390/jimaging10050099.

[11] Ceylan N, Snmez E, Kaar S. Cost effective detection of uneven mounting fault in rotary wing drone motors with a CNN based method[J].Signal, Image and Video Processing, 2024, 18(11):8049-8059.DOI:10.1007/s11760-024-03450-4.

[12] Ahmad A, Li Z, Tariq I,et al.IDSSCNN-XgBoost: Improved Dual-Stream Shallow Convolutional Neural Network Based on Extreme Gradient Boosting Algorithm for Micro Expression Recognition[J].Computers, Materials & Continua, 2025, 82(1).DOI:10.32604/cmc.2024.055768.

[13] Chen W S, Chen X L,J. LiuQ.B. WangX.F. LuY.F. Huang.Detection and recognition of UA targets with multiple sensors[J].The Aeronautical journal, 2023, 127(Feb. TN.1308):167-192.DOI:10.1017/aer.2022.50.

[14] Ashkenazi I, Benady A, Zaken S B,et al. Radiological Comparison of Canal Fill between Collared and Non-Collared Femoral Stems: A Two-Year Follow-Up after Total Hip Arthroplasty[J]. Journal of Orthopaedic Surgery and Research, 2024.