Construction and Research of Machine Learning-based Pedal Misoperation Recognition Method

Jingjun Su[a], Yunpeng Xu*

School of Information Engineering, Chang'an University, Xi'an, China.

[a]18322600137@163.com, *526835369@qq.com

Abstract. Background: A large number of traffic accidents are caused by drivers mistakenly stepping on the pedal in emergencies, highlighting an urgent concern about reducing pedal misoperation of drivers. Nowadays, machine learning has been widely applied in the fields of automatic driving, vehicle-circuit coordination, automatic obstacle avoidance, etc. However, the detection technology that focuses on the driver's driving behavior during driving has not yet been able to greatly reduce the occurrence of traffic accidents. Purpose: Data collected by Chang'an University's comprehensive driving behavior test platform were used to recognize whether a car driver has the behavior of pedal misoperation by developing and comparing a variety of machine learning algorithms. This paper proposes a method based on machine learning algorithms that takes into account the visual characteristics of the driver to identify pedal misoperation. Research methods: Five different machine learning algorithms were compared for the behavioral judgment of pedal misoperation of drivers through multiple evaluation indexes. Then, the performance of each algorithm was evaluated. As verified, the RandomForest algorithm outperforms all other algorithms with an accuracy rate of 98.4%.Conclusion: According to the research results, a method for recognizing the pedal misoperation behavior of drivers based on the RandomForest algorithm considering visual characteristics can more accurately recognize whether there is a pedal misoperation behavior.

Keywords: Machine Learning; Visual Characteristics; Pedal Misoperation; RandomForest.

1. Introduction

Statistically, 12.6% of traffic accidents each year are caused by the driver's behavior of mistakenly stepping on the accelerator pedal in emergencies[1][2]. Machine learning nowadays has been widely used in vehicle engineering, such as recognizing pedal misoperation. Hu Zhenqi et al. [3] proposed a control scheme based on the angular acceleration of throttle rotation. Lu Xiong et al.[4] constructed an ultrasonic radar-based system to avoid accelerator pedal misoperation. Mei Zhewen et al.[5] proposed a control scheme that has the accelerator pedal connective band pull the slotted ratchet. Wang Yuxing et al. [18] designed a device using electronic throttle to control pedal misoperation. Yuan Wei et al. [6] derived a pedal misoperation behavior recognition method for electric buses based on four parameters such as vehicle speed, motor torque, accelerator pedal opening and brake pedal opening, and the iForest algorithm through simulation and experiment. Zhang Tianyu et al.[19] studied the brain awareness of the driver's neurological system through medical analysis, and applied hardware logic circuits and Fuzzy PID control algorithm to build an ABS antilock safety driver assistant system. Although a wealth of research on the construction of the accelerator pedal misoperation recognition system has been conducted, the technology of how to build and adjust the machine learning algorithm for accurate and timely detection of the abnormal condition of drivers is not yet mature. Furthermore, the causes of traffic accidents are complex and varied. The detection of driving behaviors during the driving process alone cannot greatly reduce the occurrence of traffic accidents. Towards such problems, Chen Yang et al. creatively introduced the concept of "vision confusion degree" in the invention patent, "A Method and System for Recognizing Pedal Misoperation Behavior of Automobile Drivers". Drivers' behaviors combined with visual characteristics were analyzed. Based on this, a method was proposed to recognize the pedal misoperation of automobile drivers by considering the steering wheel rotational angular velocity, the driver's visual characteristics, and the distance between the vehicle and the obstacles in front[7]. However, the method is relatively thin and simple, not forming a well-established recognition system.

Compared with the various methods mentioned above, a machine learning-based behavioral recognition method that integrates the visual characteristics of drivers can perform well in determining whether the driver misoperates the accelerator pedal. Therefore, this paper used five well-known machine algorithms for comparison, including Multi Layer Perception (MLP) [8], Support vector machine (SVM) [9], TabNet[10], XGBoost[11], and RandomForest[12]. Data from on-vehicle experiments were trained and tested to evaluate the performance metrics of all algorithms and select the optimal one.

A certain amount of data collected through on-vehicle experiments constitutes a dataset, which consists of two small datasets, Dataset 1 and 2. Dataset 1 includes 3 eigenvalues and 6 target variables. In this paper, a single hidden-layer feed-forward neural network was applied to determine the driver's current facing area based on 3 eigenvalues provided in Dataset 1. Dataset 2, on the other hand, includes 4 eigenvalues and 1 target variable. The mentioned machine learning algorithm was applied to determine whether the driver misoperates the accelerator pedal based on the 4 eigenvalues provided in Dataset 2. This method is able to recognize whether the behavior of stepping on the pedal is a misoperation or not with lower cost, higher accuracy, and timeliness. In this way, it to a certain degree ensures that the vehicle's own safety guarantee system makes a timely response and ultimately protects the driver's life.

The feasibility and necessity of the method in this paper are presented in the introduction. Chapter 2 describes the materials required for and the specific path of implementation. Chapter 3 then puts forward the research process and results. Chapter 4 summarizes the analysis and conclusions regarding the research results.

2. Materials and Data Processing

2.1 Datasets

The invention patent[7] categorizes the areas that a driver may look at while driving into six ones. The dataset 1 in the research is the angular information of pitch, yaw, and roll axes of the driver's head while gazing at these six areas obtained by using an ordinary camera based on Mediapipe and a head post estimation algorithm [13]. It is an unbalanced classification dataset consisting of 3958 instances, including 3 eigenvalues and 6 target variables as shown in Table 1.

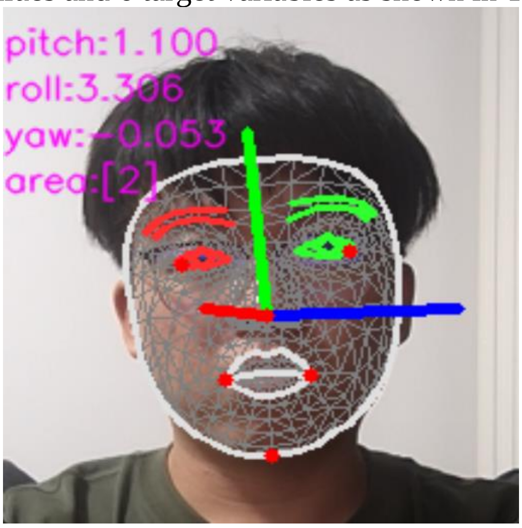

Figure 1: Specific Recognition Effect of Driver's Visual Characteristics

Table 1: 3 Eigenvalues for Dataset 1

Characteristics	Definition
pitch	Pitch axis angle of the driver's head

yaw	Yaw axis angle of the driver's head
roll	Roll axis angle of the driver's head

Dataset 2 is the data collected from Chang'an University's comprehensive driving behavior testing platform during car driving. It is an unbalanced categorical dataset consisting of 314 instances, which includes 4 eigenvalues and 1 target variable, as shown in Table 2.

Table 2: 4 Eigenvalues for Dataset 2

Characteristics	Definition
Vehicle Speed	Current speed expressed in km/h (unit)
Steering Wheel Rotational Angular Velocity	Angular speed of the steering wheel expressed in °/s (unit)
Accelerator Pedal Opening	Current opening of the accelerator pedal
Visual Confusion	Current level of visual confusion in the car driver's sight

2.2 Data Preprocessing

1) Recognizing and Dealing with Null Values

Nulls that are not properly identified and handled during data preprocessing can lead to erroneous conclusions from the data. Null values did not appear in the data of this research, so the handling of null values will not be discussed.

2) Classification and Coding of Data

Data classification refers to dividing data into different classes of information according to the need. This paper for the discussion of the driver's field of vision is a multiclassification problem that can be classified into six different classes from 0 to 5. Therefore, the categorical data is the six parts from 0 to 5 of the divided dataset. Binary 0 to 5 was used to label the dataset in the research.

In addition, the discussion of whether the pedal operation behavior is abnormal or not is a binary classification, so the categorical data is the dataset being divided into two parts, abnormal and not abnormal. Anomalies and non-anomalies were respectively labeled as "1" and "0" in this research.

3) Dividing Training Set and Test Set

The data collected were preprocessed and the entire dataset was divided into a training set and a test set, where the ratio of the training set to the test set is 80:20. The training set was used to train the model, while the test set was used to verify the degree of model training.

3. Research Methods

3.1 Multi Layer Perception

A fully connected Multi Layer Perception (MLP) neural network[8] is shown in Fig. 2.

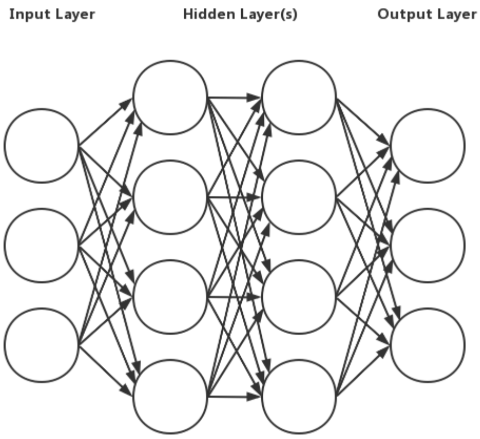

Figure 2: MLP Neural Network

This neural network consists of three layers: input layer, hidden layer and output layer. The hidden layer used to extract the eigenvalues in this network can be one layer or more. A neural network containing 6 hidden layers and 100 neurons per layer was chosen for the research. Different numbers of hidden layers and hidden neurons will lead to different results, and the choice of different activation functions will also lead to completely different results. Relu is usually used in the hidden layer as the activation function[14].

The specific equation for calculating the output of each layer of neurons is given below.

$$a^l = \sigma(W^l a^{l-1} + b^l) \tag{1}$$

a^l denotes the output of the l^{th} layer after the activation function and σ refers to the activation function. W^l represents the weight matrix of the l^{th} layer, and b^l is the bias column vector of the l^{th} layer.

This paper hopes to adjust the "connection power" between neurons and the threshold value of each functional neuron by Back Propagation (BP), that is, to realize the recognition of pedal misoperation by training the MLP model [14][15].

3.2 Support Vector Machine

Based on the dataset's eigenvalue data, Support Vector Machine (SVM) [9] can find a hyperplane with a maximum boundary in the resulting n-dimensional space to realize the algorithm for binary classification of data points. That's why the algorithm is often applied to binary classification problems.

The equation of the hyperplane with a maximum boundary is as follows.

$$\max_{\lambda_i \geq 0} W(\lambda) = \max_{\lambda_i \geq 0}\left(\sum_{i=1}^{n} \lambda_i - \frac{1}{2}\sum_{i=1,\, j=1}^{n} \lambda_i \lambda_j y_i y_j x_j^T x_i\right) \tag{2}$$

λ is lagrange multiplier. x and y are from the sample (x, y).

3.3 TabNet

TabNet[10] is a network model for tabular data that focuses on its most salient features. Therefore, it has better interpretability for more efficient learning. In classification as well as regression tasks in several fields, TabNet has shown to be no less capable than other learning models.

This research used an Attentive Transformer to get the masks, which leveraged the processed features $a[i-1]$ from the former decisions.

$$\mathrm{M}[i] = sparsemar\,(P[i-1] \cdot h_i(a[i-1])) \tag{3}$$

$M[i]$ is the feature attention assignment weight for each sample. h_i refers to a trainable function and $P[i]$ represents an a priori scaling term.

The initial tabular data obtained by TabNet does not need to be preprocessed. Similar to many other network architectures, it used gradient descent optimization for model training.

3.4 XGBoost

XGBoost[11] is a machine learning algorithm for tree-enhanced that is scalable in all cases and capable of solving large-scale problems while consuming minimal resources.

The final predicted value for the i th sample is the following.

$$\tilde{y} = \sum_{k-1}^{K} f_k(x_i), f_k \in F \tag{4}$$

x_i denotes the characteristics of the sample. $f_k(x_i)$ is the prediction result of the k th tree on the x_i sample. These values are added together to get the final result $\tilde{y}_i$

Combined with the real results y_i, the loss function can be constructed.

$$Obj = \sum_{i=1}^{n} l(y_i, \tilde{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{5}$$

l denotes the loss function and the latter term is used to control the complexity. g_i represents $\partial l(y_i, \tilde{y}_i^{<k-1>})$ and h_i denotes $\partial^2 l(y_i, \tilde{y}_i^{<k-1>})$. After derivation, the loss function only needs to be optimized as follows.

$$Obj = \sum_{i=1}^{n} [g_i \cdot f_K(x_i) + h_i \cdot f_K^2(x_i)] + \Omega(f_K) \tag{6}$$

In general, when constructing the tree, the loss function of the current tree should be calculated first. Then all the possibilities should be listed as far as possible, according to which calculates the corresponding loss function. The loss function that reduces the most is taken as the current shape of the tree, and so on continues to repeat. Until it is not reduced enough, it will be set as a threshold.

3.5 RandomForest

RandomForest[12] is an integrated learning algorithm that aims to achieve better results by integrating multiple classifiers.

The final expression of the Random Forest algorithm in the classification problem is as follows.

$$\tilde{y} = \arg\max i \sum j = 1^m I(f_j(x)) == i) \tag{7}$$

$\tilde{y}$ denotes the final classification result. $f_j(x)$ is the classification or regression result of the jth decision tree. m is the number of decision trees. In addition, I denotes the indicator function, which returns 1 when the condition is satisfied and 0 otherwise.

In this algorithm, when all the decision trees are constructed, new sample values and categories are predicted through multiple strategies. Generally, for classification problems, RandomForest inputs a new sample into all the decision trees, calculates the number of recurrences for each category, and filters the category with the highest number of occurrences as a prediction through a voting strategy.

4. Research and Results

The purpose of this research is to integrate the advantages of previous techniques and to investigate the timeliness and effectiveness of different machine learning algorithms in recognizing pedal misoperation behavior. Then, the paper proposed a method for recognizing the pedal misoperation behavior of automobile drivers based on machine learning considering the driver's visual characteristics. A single hidden-layer network model was trained to obtain real-time facial facing area information of a driver during driving. The driver's visual confusion was also calculated, which constitutes Dataset 2 with the data collected by Chang'an University's comprehensive test platform for driving behavior. Dataset 2 was used to evaluate the performance of multiple machine learning algorithms to select the best one for recognizing pedal misoperation.

4.1 Research Setup

All research was conducted in the Python platform environment[16] and the model building relied on TensorFlow[17].

4.2 Algorithm Evaluation Indexes

The ratio of data used for model training and model testing is 80:20. In this research, a set of performance indexes consisting of Accuracy, Precision, Recall, and F1Score of the models were selected. ROC curves were plotted and AUCs of the curves were computed. This set of performance indexes is based on True Positive(TP), False Positive(FP), True Negative(TN), and False Negative(FN). The performance indexes are explained below.

TP: The number of Positive instances in which the model correctly determines that the driver has pedal misoperation behavior.

TN: The number of Negative instances in which the model correctly determines that the driver does not have pedal misoperation behavior.

FP: The number of Negative instances in which the model wrongly determines that the driver has pedal misoperation behavior.

FN: The number of Positive instances in which the model wrongly determines that the driver does not have pedal misoperation behavior.

Accuracy represents the ratio of correctly judged instances to all instances in the dataset, as shown in Equation 8.

$$\mathrm{Accu}racy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

Precision refers to the ratio of correctly judged Positive instances to all judged Positive instances in the dataset, as shown in Equation 9.

$$\mathrm{Precision} = \frac{TP}{TP+FP} \tag{9}$$

Recall denotes the ratio of correctly judged Positive instances to all Positive instances in the dataset, as shown in Equation 10.

$$\mathrm{Recall} = \frac{\mathrm{TP}}{TP+FN} \tag{10}$$

F1score is the weighted average of Precision and Recall as shown in Equation 11.

$$\mathrm{F1score} = 2 \times \frac{\mathrm{Recall} \times \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}} \tag{11}$$

For the AUC-ROC curve, AUC is a model evaluation index that indicates the degree of separability. It is used as a performance measure to indicate the model's ability in classification. ROC, on the other hand, is a probability curve. The higher the AUC is, the better the model performs in recognizing pedal misoperation behavior. The ROC curve was plotted using TPR on the y-axis and FPR on the x-axis. TPR was defined in the same way as Recall. FPR was defined as follows.

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}} \tag{12}$$

4.3 Results and Discussion

Based on Dataset 1, a model that recognizes the driver's current facial facing area was trained by a single hidden layer neural network. The relationship between the recognition accuracy of the model and the number of trainings is demonstrated in Fig. 3.

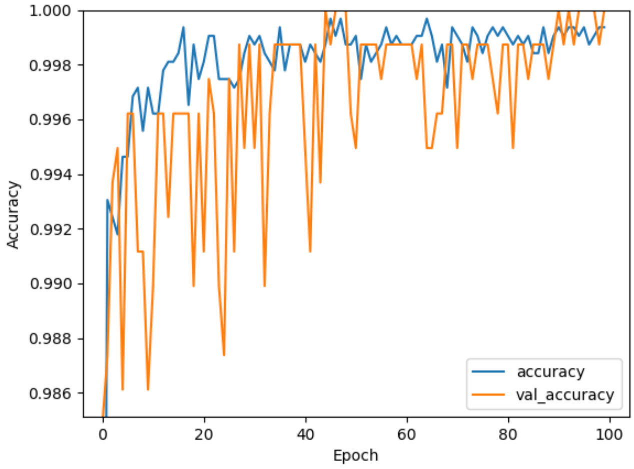

Figure 3: Relationship between Model Recognition Accuracy and Training Times

As presented in Fig. 3, the model recognition accuracy reached more than 98.6%. Through the figure and the calculation mentioned in the invention patent[7], the driver's vision confusion degree was calculated. The real-time vision confusion degree of the driver in the research is one of the features of Dataset 2.

Based on Dataset 2, this paper compared the evaluation indexes of Accuracy, AUC, Precision, Recall, and F1score algorithm of the 5 selected machine learning models after training and optimizing, as shown in Table 3.

Table 3: Comparison of Evaluation Indexes of All Models

Algorithm	Accuracy	AUC	Precision	Recall	F1score
MLP	0.937	0.881	0.933	0.778	0.848
SVM	0.911	0.825	0.923	0.667	0.774
TabNet	0.286	0.465	0.192	0.769	0.308
XGBoost	0.984	0.962	1.000	0.923	0.960
RandomForest	0.984	0.988	0.952	1.000	0.976

The curves of MLP's and SVM's ROC are as Figure 4.

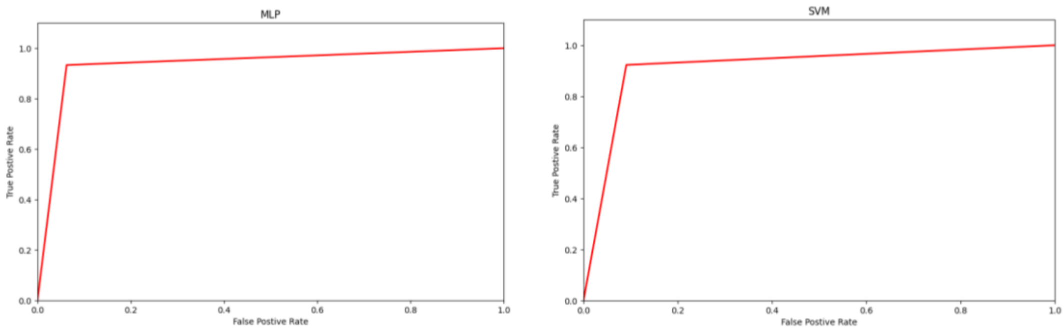

Figure 4: Curves of MLP、SVM's ROC

The curve of TabNet's and XGBoost's ROC are as Figure 5.

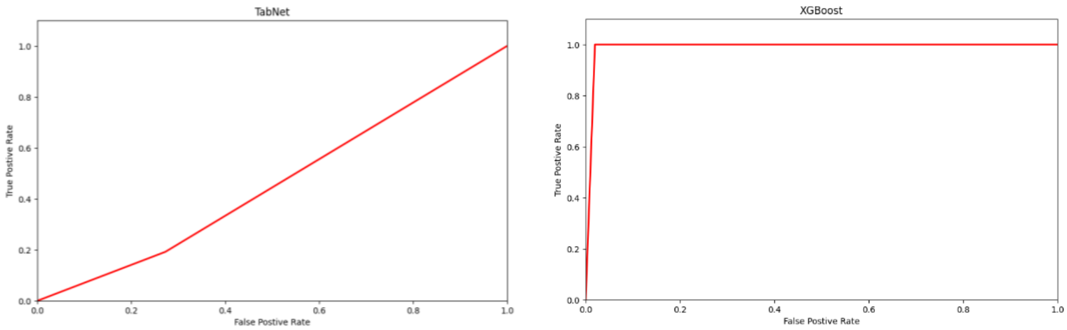

Figure 5: Curves of TabNet、XGBoost’s ROC

The curve of RandomForest’s ROC is as Figure 6.

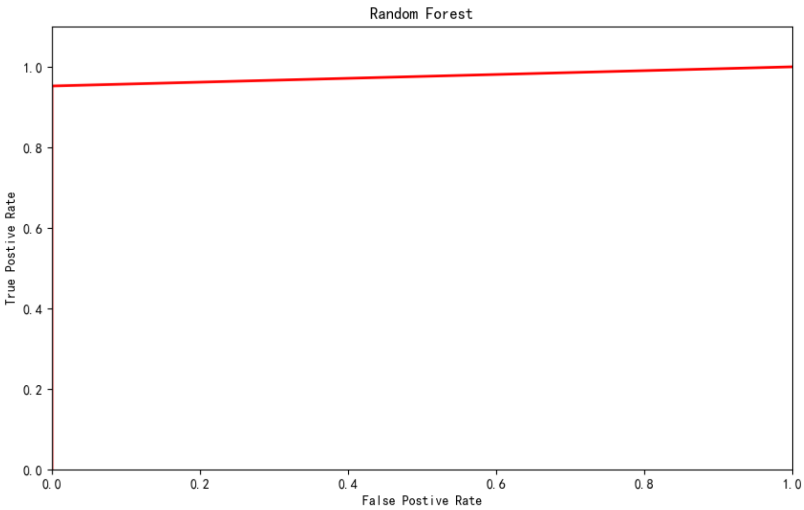

Figure 6: Curve of Random Forest's ROC

As shown in Table 3 and Figures 4 to 6, the test performance of the RandomForest model is generally better than the other models. This indicates that a proposed method for recognizing pedal misoperation based on this algorithm is the best performing, which is the most desirable solution.

5. Conclusion

With the continuous improvement of on-vehicle computing power, it is no longer out of reach to carry a machine learning model on the on-vehicle computer. The machine learning algorithm model has the ability to improve performance based on continuous self-training of data. It is undoubtedly a preferable choice for recognizing and processing solutions when considering the visual characteristics and behavioral characteristics of drivers.

The research results demonstrate that RandomForest algorithm outperforms all others, with an accuracy of 98.4%. It was tested to have the best performance in determining whether the pedal is misoperated by considering the driver’s visual characteristics. In conclusion, this recognition method based on RandomForest algorithm is promising. In the future, it is expected to further optimize and test this method in experimental platforms. More other machine learning algorithms will also be explored to compare the performance index of each and iterate the algorithms for the optimality of the model.

References

[1] ZHANG C, GONG B, HE H, et al. Special Investigation and Countermeasure Suggestions on Electric Vehicle out-of-control[C]. IOP Conf. Series: Earth and Environmental Science. IOP Publishing, 2020: 012067.

[2] LIU B.X.，LIU L.G., TIAN H.L. et al. Current Research Status of Automotive Misuse of Accelerator Pedal Solution [J]. China Foundry Equipment & Technology. 2020,55(06)

[3] HU Z.Q., ZHU C.J. LI J. et al. Development of A System for Preventing Misuse of Accelerator Pedal in Automobiles [J]. Automobile Engineering,2011,33(08):713-716.DOI:10.19562/j.chinasae.qcgc.2011.08.014.

[4] LU X., WANG Q., WEN J.F. et al. Ultrasonic Radar Based Accelerator Pedal Anti-Misstep Control System [J]. Passenger Car Technology and Research,2022,44(04):18-21.DOI:10.15917/j.cnki.1006-3331.2022.04.007.

[5] MEI Z.W., LI X.J., WANG M.H. et al. Automobile Accelerator Pedal Anti-misstep Emergency Braking Device [J]. Mechanical Design and Manufacturing,2010,No.236(10):201-203.DOI:10.19356/j.cnki.1001-3997.2010.10.079.

[6] YUAN W., YUAN X.H., GAO Y. et al. An Electric Bus Pedal Misoperation Recognition Method Based on Natural Driving Data [J]. Journal of Jilin University (Engineering Edition).2022

[7] CHEN Y., WANG C., LUO Y.W. et al. An Automobile Driver Pedal Misoperation Behavior Recognition Method and System[P]. China Patent：202310375840.5，2023.06.23

[8] Tolstikhin I , Houlsby N , Kolesnikov A ,et al.MLP-Mixer: An all-MLP Architecture for Vision[J]. 2021.DOI:10.48550/arXiv.2105.01601.

[9] Cortes C V V , Cortes C , Vapnik V ,et al.Support-vector networks[J]. Machine Learning, 20, 273-297, 1995.

[10] Sercan Ömer Arik.TabNet: Attentive Interpretable Tabular Learning[J]. CoRR. Volume abs/1908.07442 , Issue . 2019

[11] Tianqi Chen.XGBoost: A Scalable Tree Boosting System[J]. CoRR. Volume abs/1603.02754 , Issue . 2016

[12] Breiman.Random forests[J].MACH LEARN, 2001, 2001,45(1)(-):5-32.DOI:10.1023/A:1010933404324.

[13] Abate Andrea F.Head pose estimation: An extensive survey on recent techniques and applications[J].Pattern Recognition.Volume 127,Issue.2022

[14] Al Bataineh Ali.MLP-PSO Hybrid Algorithm for Heart Disease Prediction[J].Journal of personalized medicine. Volume 12 , Issue 8 . 2022. PP 1208-1208

[15] Rumelhart D E , Hinton G E , Williams R J .Learning Representations by Back Propagating Errors[J].Nature, 1986, 323(6088):533-536.DOI:10.1038/323533a0.

[16] Travis E. Oliphant.Python for Scientific Computing[J]. Computing in science & engineering. Volume 9 , Issue 3 . 2007. PP 10-20

[17] Martín Abadi,T.E.TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems[J].arXiv preprint arXiv:1907.12412,2019.

[18] WANG Y.X., LV F.Q., YANG S. et al. Solution for Automobile Accelerator Pedal Anti-misstep Device [J]. Automobile Maintenance, 2015(07):6-7.

[19] ZHANG T.Y., ZHOU Z.F., GU S.J. et al. Research on Gas Pedal Safety Assistant System [J]. Practical Automotive Technology, 2017(12):93-94. DOI:10.16638/j.cnki.1671-7988.2017.12.030.