

Study on the Correlation between Drivers' Anger Language and Driving Anger

Anmengdie Li, Jinglei Zhang *, Kaili Wang

Shandong University of Technology, Zibo Shandong 255000, China

* Corresponding author: Jinglei Zhang.

Abstract: In order to study the correlation between drivers' language and anger emotion, a questionnaire survey was conducted to collect drivers' language habits and select the anger keywords that drivers use more frequently. Then, the anger keywords were scored quantitatively from three aspects of anger degree, occurrence frequency and impact on driving safety, and the critical ratio, correlation, reliability and validity of the questionnaire were tested. The results show that the weight of "death" and "stupid" is higher than 0.1, and the weight of "God" and "drive" is less than 0.1, and the weight of "drive" is lower than 0.1. It can be seen that different anger languages have different degrees of representation effect on drivers' anger. This method can be used as one of the bases for identifying drivers' anger, which provides a theoretical basis for multi-modal drivers' anger recognition, and has certain practical significance for drivers' driving safety.

Keywords: Driving language; Driving behavior; Driving in anger, Emotion recognition; Fuzzy comprehensive evaluation.

1. Introduction

In recent years, the process of urbanization has been accelerating, the traffic pressure continues to increase, and the level of traffic environment and traffic safety is not optimistic. In the road traffic system, people are the main factor leading to traffic accidents. Affected by the driving environment, drivers will lead to angry driving behaviors, including aggressive driving behaviors such as random lane change, forced overtaking, and running red lights. These behaviors are often accompanied by the use of rude gestures, verbal insults, threats, rough behavior, and even malicious damage to others' property, which will not only affect normal traffic driving, but also seriously affect social security. Because the angry behavior of drivers is often accompanied by the output of angry language, this study aims to analyze the correlation between driver's language and anger during driving, which is of great significance to automatically identify drivers' anger through computers and ensure traffic safety.

At present, in the research on driver's angry language and anger at home and abroad, comprehensive recognition research focuses on speech features rather than semantic features.

In terms of voice features, Li Shingling et al. analyzed the spectral characteristics of the driver's voice, built the recognition model by optimized PNN (Probabilistic Neural Networks) based on Firefly Algorithm (Firefly Algorithm, FA), which has certain advantages over traditional methods. Wang wenjing used the mel frequency cepstrum coefficient (MFCC) as features to realize the road rage diagnosis system based on speech; and in noisy environment, the accuracy of anger recognition rates is 96.27% and 97.87% obtained on the RAVDESS and CASIA data sets respectively Zhang huiyun's research summarized emotional speech recognition from the aspects of speech emotion corpus, emotion feature extraction and application of speech emotion recognition, listed common emotion classification methods Nguye et al. used cascading 3-dimensional neural networks to extract the voice and face behavior features, and mixed the modes; they used the DBN algorithm to achieve feature level fusion, and the

recognition accuracy reached 90.85%. Zhang et al. used DBN fusion network to achieve model level fusion of voice information and face information, as a classification algorithm, support vector machine has a recognition accuracy of 85.97%.

In terms of language characteristics, Gao chengji proposed a language emotion recognition method, which describes the English language sentiment using the two-dimensional space model of potency-activity, and uses the staged approach based on Gaussian mixture model and support vector machine to measure the emotional valence and emotion Activation for classification and identification, with the highest accuracy of 96.9%. Chen xin proposed a metaphorical emotion identification method by compositing emotional scenes (MET-ES), this method constructs emotional representations of sentences by utilizing the emotional scenes of words, and depicts the emotional reasons by fusing emotional representations of word pairs, which are conducive to the model in order to correctly identify the emotions of sentences. Hu Qianjing studied the effect of emotion on driving behavior, established specific driving indicators significantly related to emotion from the physiological and behavioral levels, and finally proposed feasible strategies for reasonable compensation. Tim et al. conducted feature modeling in terms of language and acoustics, acoustically, statistical data are generated according to acoustic audio descriptors; linguistically, word and phrase models based on probability and entropy are used to analyze their applications in anger classification. Gaonka et al. designed tag semantics to better encode semantic information related to emotion. In order to solve the multi label problem of text emotion, Li et al. established a factor graph based on emotion labels and upper and lower instance dependencies, and designed a factor graph reasoning model. In order to solve the problem of insufficient annotation data in emotion recognition tasks and improve the results of resource poor Chinese emotion recognition, Zhang et al. used resource rich English corpora to construct English text emotion recognition tasks to help Chinese text emotion recognition.

To sum up, anger language research focuses on speech

recognition, using typical speech feature parameters such as Mel-Frequency Cepstral Coefficients (MFCC), Linear Prediction Coefficients (LPC), Linear Predictive Cepstral Coefficients (LPCC) to model and complete anger emotion recognition; model performance is improved on the basis of face recognition, to prove the correlation between voice and emotion recognition. As for language research, domestic research is not extensive, and most of them prefer theoretical analysis of language texts. In order to further analyze the correlation between drivers' angry language and anger, the questionnaire survey method is used to select the typical anger language in drivers' angry driving, establish the rating system of anger words, explore the correlation between language words and anger, and provide theoretical support for the recognition of multimodal drivers' anger.

2. Questionnaire investigation

The text designed a questionnaire for two purposes. First, it is necessary to understand the driver's language habits and establish an angry language database so that angry language can be selected and graded through the result statistics; second, it is necessary to investigate and count which driving behaviors are more likely to irritate drivers during driving, so as to more effectively induce drivers to generate anger during the simulation driving experiment. Therefore, the questionnaire includes anger language collection questionnaire and anger language rating questionnaire.

2.1. Angry language collection questionnaire

In order to establish an angry language database that can accommodate a certain sample size, this paper designs a subjective driver anger language collection table, which includes the basic information of the driver, such as driver's gender, age, driving age, whether the driver can express anger in words during driving, and what angry language he can say.

The questionnaire is distributed by electronic questionnaire using multi-stage stratified overall sampling method to ensure the reliability of the data. Table 1 shows the statistical results of the basic information of drivers. It can be seen that the proportion of men and women in the sample is balanced, and the age group is mainly young, which is in line with the current trend of young drivers. The population with different driving ages is relatively balanced, reflecting the scientific and representative nature of this survey.

Table 1. Statistics of driver's basic information

Features	Classification	Proportion
Gender	Male	49.95%
	Female	50.05%
Age	18-25 years old	28.35%
	26-35 years old	21.62%
	36-45 years old	20.81%
	45-55 years old	19.51%
	56 years old and above	9.71%
Driving age	2 years and below	20%
	3-5 years	35.86%
	6-10 years	25.52%
	more than 10 years	18.62%

A total of 743 angry words were collected in the questionnaire, through word frequency analysis of the text, 420 angry words were obtained, such as "angry death" and "sick". Among them, 10 words appeared more than 20 times, and the rest appeared less than 8 times. There is a large gap in

the data frequency, so the top 10 words of the frequency are selected as angry language keywords. Since the frequency of anger language keywords used among drivers can reflect the language habit characteristics of drivers when they are angry to some extent, the higher the frequency of use, the better the driver's anger can be represented. Therefore, this paper continues to use these 10 keywords for the rating of anger language. The complete spelling and frequency of the main keywords are shown in Fig. 1.

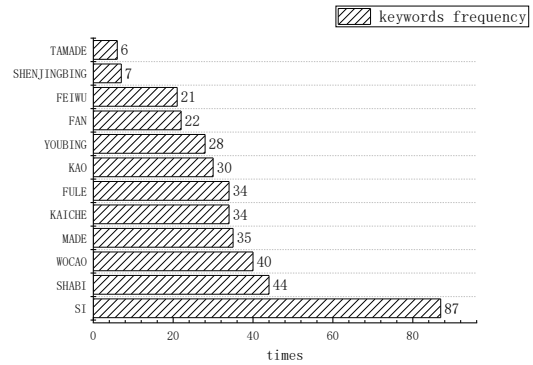


Fig. 1 Figure of Angry Word Frequency Statistics

2.2. Angry language rating questionnaire

This survey imitates the Likert 5 scoring method so that it can quantitatively evaluate keywords. Evaluate the 10 anger keywords from three aspects: the degree of anger ("extremely light", "slightly light", "average", "slightly heavy", "extremely heavy"), the frequency of occurrence ("never", "occasionally", "generally", "often", "always"), and the impact on driving safety ("extremely light", "slightly light", "average", "slightly heavy", "extremely heavy"), and assign the options one to five points in order to quantify the evaluation. A total of 642 questionnaires were distributed in the second stage, with an effective questionnaire rate of 96.1%. Table 2 shows some examples of questionnaires.

Table 2. Some examples of questionnaires

1. If you say "death" in anger during driving, your anger level is
A. extremely light (Score 1 point)
B. slightly light (Score 2 point)
C. average (Score 3 point)
D. slightly heavy (Score 4 point)
E. extremely heavy (Score 5 point)
2. During driving, you think the frequency of angry words "death" is
A. never (Score 1 point)
B. occasionally (Score 2 point)
C. generally (Score 3 point)
D. often (Score 4 point)
E. always (Score 5 point)
3. In the process of driving, do you think the degree of influence of saying "death" in anger on driving safety is
A. extremely light (Score 1 point)
B. slightly light (Score 2 point)
C. average (Score 3 point)
D. lightly heavy (Score 4 point)
E. extremely heavy (Score 5 point)

2.2.1. Critical ratio analysis

The total scores of the survey scale are arranged in descending order, and the first 30% is selected as the high group, and the last 30% is selected as the low group. The difference of each item between the high and low scores of the scale is tested by the method of independent sample t-test to obtain the critical ratio (CR) of the scale. The calculation

formula is as follows

$$CR = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \delta_i} \quad (1)$$

In formula (1), c is the factor load of each item, δ_i is the residual variance of each item. The results showed that the difference between the two groups was significant, the CR value was less than 0.04, and the item discrimination ability was good.

2.2.2. Critical ratio analysis

Table 3. Distribution table of correlation coefficient between each item and category grouping of the survey scale

Project	Category grouping		
	Anger level group	Frequency group	Impact degree group
Anger level group	1.000	0.139	0.045
Question1	0.574	0.142	-0.123
Question 4	0.732	0.182	0.012
Question 7	0.665	0.114	0.041
Question 10	0.589	0.132	0.062
Question 13	0.884	0.122	0.013
Question 16	0.843	0.113	0.011
Question 19	0.744	0.143	0.021
Question 22	0.789	0.110	0.025
Question 25	0.765	0.120	0.029
Question 28	0.777	0.161	0.027
Frequency group	0.139	1.000	0.214
Question 2	0.155	0.772	0.332
Question 5	0.145	0.735	0.374
Question 8	0.120	0.659	0.297
Question 11	0.114	0.689	0.325
Question 14	0.132	0.894	0.359
Question 17	0.123	0.742	0.319
Question 20	0.143	0.721	0.298
Question 23	0.144	0.855	0.308
Question 26	0.162	0.746	0.364
Question 29	0.102	0.765	0.337
Impact degree group	0.045	0.214	1.000
Question 3	0.012	0.288	0.772
Question 6	-0.089	0.292	0.715
Question 9	0.033	0.304	0.743
Question 12	0.045	0.335	0.789
Question 15	0.043	0.342	0.776
Question 18	0.025	0.299	0.751
Question 21	-0.016	0.332	0.699
Question 24	0.012	0.300	0.803
Question 27	0.036	0.315	0.798
Question 30	0.034	0.314	0.810

The questionnaire items were classified into anger degree group, frequency group and influence degree group according to the content of the questions, and the correlation between

the three category item groups and between each small item question and category group was explored. The correlation coefficient between each item and category item group of the survey scale is shown in Table 3. It can be seen from the table that the correlation coefficient between the items of the scale is relatively significant, especially between the frequency group and the impact degree group, which indicates that the internal consistency between the items is good, and can provide satisfaction and reliability levels with good quality.

2.2.3. Reliability analysis

The text adopts Cronback's α Coefficient to measure the reliability to analyze the reliability and effectiveness of the questionnaire the formula is [18]

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum S_i^2}{S^2} \right) \quad (2)$$

In formula (2), K indicates the number of items in the question, S_i^2 is the variance within the question of the score of question i , S^2 is the variance of the total score of all questions.

SPSS21.0 software was used to test the reliability of anger language rating questionnaire. There were 30 questions in the questionnaire. After excluding invalid questionnaires, there were 617 rows of valid data. Overall Cronback's α the coefficient is 0.732, and the reliability of the two retests is 0.722 and 0.707 respectively. It can be considered that the scale is highly reliable, Cronback's of other dimensions α the coefficient and retest reliability are above 0.70, which indicates that the reliability of the scale is good. The inspection results are shown in Table 4.

Table 4. Reliability test results

	Anger level	Frequency of occurrence	Impact degree	Overall scale
Cronback's α	0.782	0.797	0.819	0.732
Retest Reliability 1	0.701	0.705	0.703	0.722
Retest Reliability 2	0.723	0.710	0.709	0.707
Number of items	10	10	10	30

2.2.4. Validity analysis

In this paper, we need to analyze the validity of the survey scale through structural validity analysis. The results showed that Bartlett's spherical test χ^2 value was 5721.099 ($P=0.000<0.01$), indicating that it was suitable for factor analysis; KMO value is 0.799, indicating that factor analysis is effective. The mode matrix, eigenvalue and variance interpretation rotated by the maximum variance method are extracted by principal component analysis, and three common factors with eigenvalue greater than 1 are selected, as shown in the simplified gravel diagram in Fig.2. The abscissa is 30 items, and the ordinate is its corresponding eigenvalue. It can be seen from the figure that the eigenvalue of three factors is significantly higher than 1, and the cumulative variance contribution rate of three common factors is 70.85%. According to the contents of items included in each factor, three factors are determined as anger degree, frequency of occurrence and impact on driving safety. The factor loads of each project are shown in Table 5.

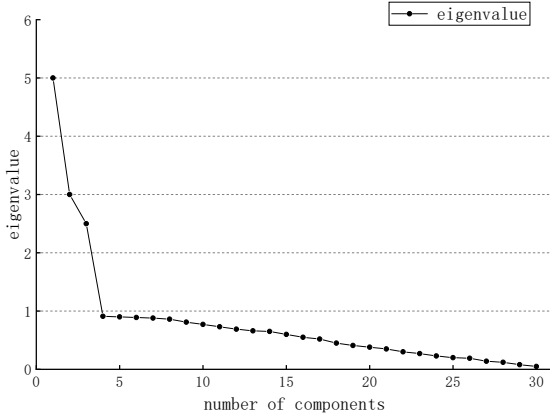


Fig. 2 Simplified rubble diagram

Table 5. Factor Load Table of Each Project

Anger level		Frequency of occurrence		Impact degree	
project	load	project	load	project	load
Question 1	0.611	Question 2	0.614	Question 3	0.872
Question 4	0.719	Question 5	0.672	Question 6	0.817
Question 7	0.650	Question 8	0.633	Question 9	0.684
Question 10	0.798	Question 11	0.816	Question 12	0.680
Question 13	0.863	Question 14	0.685	Question 15	0.870
Question 16	0.632	Question 17	0.635	Question 18	0.681
Question 19	0.533	Question 20	0.899	Question 21	0.802
Question 22	0.655	Question 23	0.702	Question 24	0.877
Question 25	0.810	Question 26	0.709	Question 27	0.676
Question 28	0.663	Question 29	0.898	Question 30	0.725

3. Evaluation analysis

3.1. Fuzzy comprehensive evaluation

Since the judgment of anger is subjective, fuzzy and difficult to quantify, the fuzzy comprehensive evaluation method is selected for analysis, it is characterized by clear and systematic results. Qualitative evaluation can be converted into quantitative evaluation according to the membership theory of fuzzy mathematics, that is, fuzzy mathematics can be used to make an overall evaluation of things or objects restricted by multiple factors. The research ideas of fuzzy comprehensive evaluation are as follows:

(1) Determine the indicator set of the evaluation object, namely the factor set. Suppose there are n evaluation indicators, which can be expressed as:

$$U = \{u_1, u_2, \dots, u_n\} \quad (3)$$

In formula (3), $u_i (i=1, 2, \dots, n)$ is the evaluation index.

(2) Determine the evaluation result set of the evaluation object, that is, the evaluation set. Suppose there are m evaluation grades, which can be expressed as:

$$V = \{v_1, v_2, \dots, v_m\} \quad (4)$$

(3) Determine the weight set of evaluation indicators, that is, the weight set. Suppose there are n evaluation indicators,

which can be expressed as:

$$A = \{a_1, a_2, \dots, a_n\}, \quad a_i > 0, \quad \sum_{i=1}^n a_i = 1 \quad (5)$$

(4) Determine the membership function and calculate its membership degree, and form the factor evaluation matrix R , that is to say, suppose the evaluation of the i th index u_i :

$$R_i = \{r_{i1}, r_{i2}, \dots, r_{im}\} \quad (6)$$

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{bmatrix} \quad (7)$$

In formula (6), $R_i (i=1, 2, \dots, n)$ is a single factor fuzzy subset, In formula (7), R is the evaluation matrix on V . (5) multiplies the factor evaluation matrix by the weight set to obtain the comprehensive evaluation model:

$$B = A \circ R = (b_1, b_2, \dots, b_n) \quad (8)$$

3.2. Entropy weight method

The entropy weight method is often used to determine the weight set if there is no authoritative data standard, it uses information entropy to calculate the entropy weight of each index according to the variation degree of each index, and gets more objective index weight after correction. The research idea of entropy weight method is as follows:

(1) Determine indicator data matrix:

If different indicator types are different, indicator forward transformation is required, that is, all indicators are unified into smaller and better indicators or larger and better indicators. Assuming that there are m projects to be evaluated, a total of n evaluation indicators are set, which can form an indicator data matrix:

$$X = (x_{ij})_{m \times n} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (9)$$

(2) Data standardization

In order to unify the dimensions of each evaluation index, the data needs to be standardized, in the standardization method, the maximum minimum standardization method is commonly used, and its calculation formula is:

$$X' = z_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} = \frac{x_{ij} - \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}}{\max\{x_{1j}, x_{2j}, \dots, x_{nj}\} - \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}} \quad (10)$$

(3) Calculating information entropy

First, calculate the probability matrix P , that is, the proportion of the i th sample under index j , and the calculation formula is:

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^n z_{ij}} \quad (11)$$

Then calculate the information entropy corresponding to each index, and normalize it to obtain its respective entropy weight, the calculation formula is:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln(p_{ij}), j = 1, 2, \dots, m \quad (12)$$

$$d_j = 1 - e_j \quad (13)$$

$$W_j = \frac{d_j}{\sum_{j=1}^m d_j} \quad (14)$$

The formula (12) represents the size of information entropy, the greater the information entropy, the smaller the

corresponding amount of information; in Formula 13, d_j is the value of information utility; in Formula 14, the entropy weight W_j of each index is the result of normalization of information utility value.

(4) Empowerment

To give weight to the data, calculate the final score, the calculation formula is:

$$S_i = \sum_{j=1}^m W_j z_{ij} \quad (15)$$

4. Result analysis

In order to conduct quantitative rating of anger for 10 anger keywords in the questionnaire, the results were analyzed according to the steps of fuzzy comprehensive evaluation.

The evaluation indicators of the object include the degree of anger of the key words, the frequency of occurrence and the degree of impact on driving safety, so the factor set is taken as follows:

U= {the degree of anger, the frequency of occurrence, the degree of impact on driving safety}

The evaluation result is 10 angry keywords, so the evaluation set is taken:

V={dead, stupid, Dammit, fuck, driving, taking, cor, sick, annoying, waste }

Using entropy weight method to determine the weight of factor set, the value of m is 10, the value of n is 3, according to the entropy value calculation formula (14):

$$W_1 = 0.1086, W_2 = 0.0979, W_3 = 0.1163, W_4 = 0.1103, W_5 = 0.0876, \\ W_6 = 0.1121, W_7 = 0.1025, W_8 = 0.1150, W_9 = 0.0839, W_{10} = 0.0658$$

According to formula (15), $S_1 = 0.4766, S_2 = 0.2317, S_3 = 0.2917$

Therefore, the weights of the three indicators are S_1, S_2 and S_3 respectively, and the weight set is determined:

$$A = \{0.4766, 0.2317, 0.2917\}$$

The membership matrix of the three factors is:

The degree of anger R1 equals:

$$(0.1111, 0.1245, 0.0895, 0.0985, 0.0798, 0.0864, 0.1004, 0.1145, 0.0921, 0.1032)$$

The frequency of occurrence R2 equals:

$$(0.1123, 0.1233, 0.0907, 0.0973, 0.0810, 0.0852, 0.1016, 0.1133, 0.0933, 0.1020)$$

The degree of impact on driving safety R3 equals:

$$(0.0866, 0.1000, 0.1140, 0.1230, 0.1122, 0.1188, 0.0680, 0.0821, 0.0965, 0.0988)$$

R1、R2、R3 Evaluation matrix of three fuzzy subsets:

$$R = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} = \begin{bmatrix} 0.1111 & 0.1245 & 0.0895 & 0.0985 & 0.0798 & 0.0864 & 0.1004 & 0.1145 & 0.0921 & 0.1032 \\ 0.1123 & 0.1233 & 0.0907 & 0.0973 & 0.0810 & 0.0852 & 0.1016 & 0.1133 & 0.0933 & 0.1020 \\ 0.0866 & 0.1000 & 0.1140 & 0.1230 & 0.1122 & 0.1188 & 0.0680 & 0.0821 & 0.0965 & 0.0988 \end{bmatrix}$$

The comprehensive evaluation model is obtained by multiplying the factor evaluation matrix by the weight set:

$$B = A \circ R = (0.1042, 0.1171, 0.0969, 0.1054, 0.0895, 0.0956, 0.0912, 0.1048, 0.0937, 0.1016)$$

According to the model results, the weight distribution of 10 keywords is shown in Figure 3, it can be seen intuitively in the figure that the weight of "dead", "stupid", "fuck", "sick" and "waste" is higher than 0.1, so these five keywords are listed as Level I angry words; the weight of "Dammit", "driving", "surrender", "cor" and "annoying" is less than 0.1, which is listed as Level II angry words. In the following research, anger language can be used as a feature vector for angry driving recognition system according to the

classification.

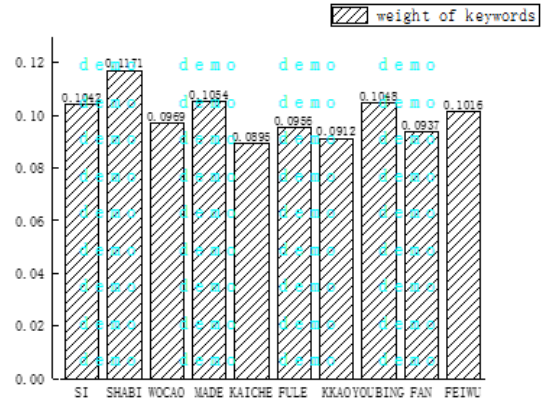


Fig. 3 Anger keyword rating weight analysis chart

5. Conclusion

This paper collects anger keywords commonly used by drivers through questionnaire survey, quantifies and scores anger keywords from three aspects of anger degree, frequency and impact on driving safety, determines the weight of three indicators using entropy weight method, and empowers anger keywords through fuzzy comprehensive evaluation. The conclusions are as follows:

(1) The anger language rating questionnaire designed from three aspects of anger degree, frequency and impact on driving safety has high reliability and validity, which can be used as a reference for anger language collection.

(2) Five first level anger words, such as "death" and "stupid", and five second level anger words, such as "Dammit" and "driving", were selected. It can be seen that different angry words and sentences have different degrees of representation effects on drivers' anger. This method can be used as one of the bases for the computer to automatically identify drivers' anger.

References

- [1] Shangqing Li, Xiaoyuan Wang, Yang Zhang, Hao Li, Hui Xiang, "Road rage recognition based on improved MFCC fusion feature and FA-PNN", Computer Engineering and Application, 2022, Vol. 5(12), p1-9.
- [2] Wenjing Wang, "Road Rage Recognition System Based on Speech Features", Zhejiang University, 2022, Vol. (6), p1-20.
- [3] Huiyun Zhang, "An Overview of Speech Recognition Systems", Information Communications, 2019, Vol. (11), p58-60.
- [4] NGUYEN D, NGUYEN K, SRIDHARAN S, et al, "Deep spatio-temporal feature fusion with compact bilinear pooling for multimodal emotion recognition", Computer vision and image understanding, 2018, Vol. 17(4), p33-40.
- [5] ZHANG S, ZHANG S, HUANG T, et al, "Learning affective features with a hybrid deep model for audio-visual emotion recognition", IEEE transactions on circuits and systems for video technology, 2017, Vol. 28(10), p3030-3043.
- [6] Chengji Gao, "A Study of Language Emotion Recognition", Computer & Digital Engineering, 2019, Vol. 47(07), p1622-1626.
- [7] Xin Chen, "Language Device and Implicit Emotion Identification Based on Language Granularity", Shanxi University, 2021, Vol. (1).

- [8] HU qianjing, "Research on Impact of Anger-Neutral Emotion on Driving Behavior in Typical Scene", Chongqing University, 2019, Vol. (1).
- [9] Tim Polzehl, Alexander Schmitt, Florian Metze, Michael Wagner, "Anger recognition in speech using acoustic and linguistic cues", *Speech Communication*, 2011, Vol. 53(9-10), p1198-1209.
- [10] Radhika Gaonkar, Heeyoung Kwon, Mohaddeseh Bastan, Niranjan Balasubramanian, Nathanael Chambers, "Modeling label semantics for predicting emotional reactions ", in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, p4687-4692.
- [11] Shoushan Li, Lei Huang, Rong Wang, Guodong Zhou, "Sentence-level emotion classification with label and context dependence", in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 2015, p1045-1053.
- [12] Lu Zhang, Liangqing Wu, Shoushan Li, Zhongqing Wang, Guodong Zhou, "Cross-lingual emotion classification with auxiliary and attention neural networks", in *Proceedings of the CCF International Conference on Natural Language Processing and Chinese Computing*, 2018, p429-441.
- [13] Lei Zhang, Jinshuan Peng, Xiaoli Chen, "Design of driving distraction behavior scale and analysis of influencing factors", *Chinese Safety Science Journal*, 2021, Vol. 31(11), p39-46.
- [14] Huihui Zhu, Rui Li, "Research on satisfaction analysis and optimization strategy of urban public transport service based on Likert scale", *Journal of Changsha Social Work College*, 2020, Vol. 27(02), p72-76.
- [15] Lu Zhang, Yun-te-na Wu, Tonglin Jin, Yanru Jia, "The Development of the Online Prosocial Behavior Scale and the Test of Reliability and Validity among College Students", *Chinese Journal of Applied Psychology*, 2022, Vol. 28(03), p270-276.
- [16] Hu Lei, Chaozhong Wu, Hui Zhang, "A Study of Driving Anger Expression Scale Based on Driving Behavior", *Journal of Wuhan University of Technology (Transportation Science and Engineering)*, 2012, Vol. 36(04), p836-839.
- [17] Hu Zhang, Maofeng Tian, "Application of reliability analysis in questionnaire design", *Statistics and decision-making*, 2007, Vol. (21), p25-27.
- [18] Xingbing Bu, Hongji Wu, Qiang Feng, Chong Liao, Yuanchao Feng, Huizhen Wang, "Analysis of main sources of atmospheric pollutants based on fuzzy comprehensive evaluation", *Environmental Protection and Technology*, 2022, Vol. 28(03), p8-13.
- [19] Huan Liu, Zhe Kai, "Research on the Evaluation of the Satisfaction of SMEs' Business Environment in Jiangbei New Area, Nanjing--Based on AHP entropy weight method", *Operation and management*, 2022, Vol. (08), p109-115.
- [20] Xinglong Wang, Ziyang Chen, Yan Liu, "Research on monitoring information quality of ATC based on combination weighting ", *China Safety Science Journal*, 2022, Vol. 32(1), p41-50.