IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 February 2024, accepted 18 March 2024, date of publication 22 March 2024, date of current version 9 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3380825

RESEARCH ARTICLE

A Takeover Risk Assessment Approach Based on an Improved ANP-XGBoost Algorithm for Human–Machine Driven Vehicles

TAO WANG[1], YAXI HAN[1], WENYONG LI[1], XIAOFEI YE[2], AND QUAN YUAN[3]
[1]Guangxi Key Laboratory of Intelligent Transportation System, Guilin University of Electronic Technology, Guilin 541004, China
[2]Faculty of Maritime and Transportation, Ningbo University, Ningbo 315211, China
[3]State Key Laboratory of Automotive Safety and Energy, School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

Corresponding author: Tao Wang (wangtao@guet.edu.cn)

This work was supported in part by National Natural Science Foundation of China [grant number 52262047, 52072214], in part by Guilin Key R&D Program [grant number 20210214-1], Liuzhou Key R&D Program [grant number 2022AAA0103], and Innovation Project of Guilin University of Electronic Technology Graduate Education [grant number 2022YCXS225].

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Review Committee of the School of Architecture and Transportation Engineering, Guilin University of Electronic Technology under Approval No. 2021PS03, and performed in line with the Helsinki Declaration.

ABSTRACT This study evaluates the risk level of human-machine collaborative driving takeover in a highway environment under the interaction of non-driving related tasks and takeover request prompt scenarios. Using a driving simulator, a 5×5 factor analysis examined non-driving tasks and takeover prompts. Takeover impact and risk indicator features were extracted, with risk indicators weighted using an improved ANP method. On this basis, the XGBoost algorithm was employed to identify crucial variables that reflect the level of takeover risk and to construct a driver takeover risk assessment model. The evaluation results indicate that the risk indicator feature with the greatest influence on the takeover risk level was the minimum TTC, which had the highest correlation ($r=-0.81$); the scene factor in the takeover influence feature had the highest correlation with the takeover risk level ($r=-0.78$), which had the greatest influence on driver takeover safety; Although non-driving related tasks exhibited a weak correlation with takeover reaction time, steering reaction time, and minimum TTC, the effect was minor. The XGBoost algorithm-based risk assessment model demonstrated superior performance over LightGBM and SVM, with 87.1% accuracy. Overall, this study highlights the significant influence of takeover scenes and minimum TTC on collaborative driving risk, enabling accurate risk modeling.

INDEX TERMS Human–machine driven vehicles, improved ANP-XGBoost algorithm, takeover risk assessment.

I. INTRODUCTION

In SAE Level 3 vehicles in autonomous driving mode, the system is capable of performing all dynamic driving tasks independently, thereby freeing the driver to engage in activities such as watching videos or playing games without needing to attend to the steering wheel and pedals [1], [2]. However, given the current stage of autonomous vehicle development, it is necessary to consider the risks arising from both the failure of the autonomous driving system and the unintended consequences of human error and environmental interference in situations where the system may not be sufficiently functional for the given environment, which is referred to as Safety Of The Intended Functionality (SOTIF) [3], [4]. In the event of a complex situation beyond the control domain, the driver is required to takeover control of the vehicle, which may lead to an accident if the driver is unable to respond in time or ensure vehicle stability after taking over [5], [6]. As such, it is necessary to conduct an in-depth study of the factors and

The associate editor coordinating the review of this manuscript and approving it for publication was Lorenzo Ciani.

interaction mechanisms that influence the level of takeover risk [7].

In previous studies, the impact of non-driving related tasks (NDRTs) on the quality of takeover has been extensively explored [8], [9]. Wu et al. [10] investigated the effect of different non-driving-related tasks and different takeover scenarios on takeover quality in a highway, and the experimental results showed that the effect of non-driving-related tasks on driving load was more significant than takeover scenarios. Shi and Bengler [11] found that the similarity of NDRTs to driving tasks can affect the takeover time, with similar tasks resulting in shorter takeover times. For instance, playing Tetris results in less takeover time and less maximum longitudinal acceleration than watching a documentary or reading and typing a summary.

Other scholars have also studied the effect of takeover request mode and takeover request time on takeover quality [12], [13]. Bazilinskyy et al. [14] investigated and found that in high urgency situations, multimodal takeover requests (TORs) were the most popular choice. In contrast, auditory TORs were preferred in low urgency situations. Tan and Zhang [15] conducted a network-based supervised experiment and showed that longer TOR lead times had a positive effect on driver situational awareness to regain control and exit the highway, with the effect plateauing at lead times of 16-30 s. Xu et al. [16] investigated driver takeover performance and workload at different levels of automation, time budgets, and road curvature. Their results showed that drivers in level 3 takeovers performed worse and had higher workloads than level 2 takeovers when entering curves with limited time budgets. Shi et al. [17] conducted driving simulation trials and found that shorter request times (6 s to 3 s) and larger secondary task loads (audio-visual combinations) significantly reduced takeover performance. Kaye et al. [18] analysed driver reaction time and vehicle lateral position and found that the use of a handheld mobile phone under human-machine cooperative driving conditions did not negatively affect the quality of vehicle takeover.

The impact of personal characteristics on takeover performance has also received considerable attention. Li et al. [19] found that females showed a smaller proportion of emergency takeovers, faster reaction times, and more stable steering operations compared to males. However, Loeb et al. [20] reached the opposite conclusion based on experimental data from a driving simulation trial of 60 drivers, where they found that males had fewer crashes (38%) than females (43%). Several studies have concluded that older people take longer to take over vehicles than younger people [21], [22], [23], [24]. On the other hand, Zhang et al. [25] conducted a meta-analysis of 129 studies and found no consistent evidence of age's effect on takeover quality. Chen et al. [26] designed a driving simulation test to compare the takeover performance of novice and experienced drivers. The study results showed that novice drivers were less stable and less adaptable in taking over, although driving experience did not significantly affect takeover time or minimum crash time.

There has not been focused research on the evaluation of the quality of human-machine collaborative driving takeover. Some scholars have used takeover quality indicators to assess the takeover risk. Gold et al. [27] investigated the takeover performance in the takeover scenarios of Level 3 conditional autonomous driving and proposed a regression model with four takeover indicators, including takeover time, shortest time to collision (TTC), braking application, and probability of collision. Happee et al. [28] investigated the aftereffects of autonomous driving in takeover scenarios. It was found that the minimum time to collision (TTC) and the clearance towards the obstacle are complimentary surrogate safety metrics for obstacle avoidance behaviour. Katharina Wiedemann et al. [29] evaluated the quality of the takeover using several metrics of the takeover time, the takeover of the lateral control (e.g., the standard deviation of the steering wheel angle) and the longitudinal control (e.g., the standard deviation of the speed). Wu et al. [30] investigated the effect of different emergency situations and takeover request lead times on takeover performance and safety. They found that within a certain range, longer takeover request lead times produced shorter takeover response times, but the surge in driving load causes a reduction in takeover safety. Xian et al. [31] developed a driving subtask safety evaluation model based on network analysis and used the evaluation model using network analysis and found that subtask driving, sub-task type and driving experience can affect driving safety. Lin et al. [32] conducted a takeover scenario test with various subtasks and constructed a takeover safety evaluation model using binary logistics regression. The evaluation test results showed that subtasks increase the takeover reaction time and reduce the safety of takeover to some extent but have no effect on the stability of takeover. However, understanding the relationship between certain factors and takeover performance is not enough to assess predicted driver takeover performance in the real world since many influencing factors may inherently interact with each other.

The current research on the direction of man-machine co-driving takeover is mostly aimed at the analysis of the factors affecting the takeover quality, and there are fewer studies on the takeover risk assessment, and most of the studies on the assessment of automatic takeover use a single subjective or objective assessment method, and lack of a comprehensive assessment of the indicators. Most of the existing studies use statistical models, which are usually subject to strict assumptions during the data processing phase, and failure to follow the assumptions may lead to unreasonable evaluation results. Compared with statistical methods, machine learning is more capable of handling large-scale and high-dimensional data, and can dig deeper into the interactions between factors for comprehensive risk assessment [33].

This paper aims to develop a risk assessment model that highlights the importance of subjective risk indicator features while reflecting objective takeover data information, and to evaluate the level of takeover risk associated with each takeover prompt scenario when performing different NDRTs. Firstly, we have designed a test featuring a gradient decrease in takeover prompting effectiveness for different NDRTs. Based on this, we developed a takeover risk assessment model that integrates subjective and objective evaluations through an improved ANP method and an integrated learning algorithm. This paper explores the characteristic variables that affect driver takeover risk and the degree of influence, providing a scientific basis and theoretical support for predicting the level of driver takeover risk and enhancing the safety of autonomous driving. The assessment framework is depicted in Figure 1.

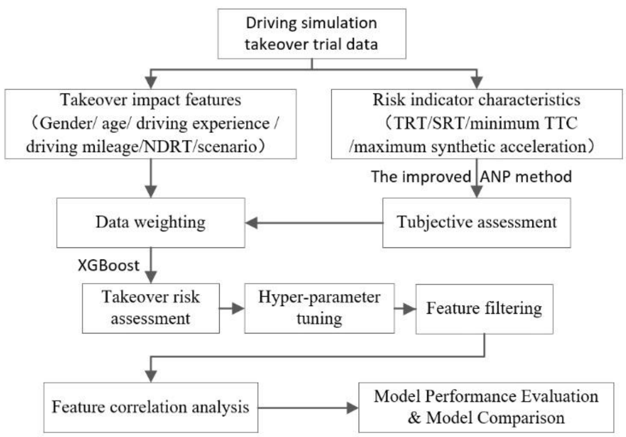

FIGURE 1. Assessment framework.

II. MATERIALS AND METHODS

A. EXPERIMENTAL DESIGN

1) THE NON-DRIVING RELATED TASKS

Engaging in NDRTs causes the driver to be unable to regain baseline ability to react and control immediately when faced with a hazardous situation [34]. To investigate the impact of realistic NDRTs on takeover risk, the trial selected four typical task categories for testing: listening to the radio, playing games, watching videos and listening to the radio + playing games (Table 1). During the automated driving phase, participants who engaged in visual or audiovisual loads were not permitted to observe the driving environment. Using chi-square test evaluation, there were no significant differences in the distribution of gender(p=0.987), age(p=0.641), driving experience(p=0.742), and driving mileage(p=0.450) among participants in the five task groups.

TABLE 1. The NDRT.

#	NDRTs	Description & Requirements
1	no task	In both automated and manual driving phases, the driver has to monitor the driving environment at all times.
2	listening to the radio	Auditory load - Music and news radio broadcasts were played throughout.
3	playing games	Visual load - Driver playing mobile game "Animal Restaurant".
4	watching videos	Audiovisual load - Driver watching a laptop showing the documentary "Tongue Tied" in the right front.
5	listening to the radio and playing games	Audio-visual load - radio broadcast while playing mobile phone mini-games.

2) TAKEOVER SCENARIOS

The test drive was on a 45km long two-way 6-lane motorway, each lane 3.75m wide. The speed limit of the road was 120km/h. The traffic flow was set to a thin traffic flow of 300 vehicles/h in one direction, operating at an average speed of 100km/h. To prevent weather conditions from affecting the test, clear, dry, and windless weather conditions were selected. The scenario intervals and road lengths were varied to avoid too regular intervals between scenarios.

To investigate changes in takeover risk level for scenarios where the urgency rises and the prompting effectiveness gradient decreases, the takeover request prompt was designed for 6s, 5s, 4s, 3s and no prompt [35]. In the 6s, 5s, 4s and 3s request time scenarios, a broken-down vehicle with a speed of 0 km/h appears in front of the vehicle, accompanied by a takeover request message: auditory (‘‘ding-ding’’ sound) and visual (‘‘Please be aware of the takeover vehicle’’). As vehicle speeds on highways are faster than on city roads, SOTIF problems are more likely to occur when the system cannot detect obstacles in time. In the SOTIF scenario, a wild animal suddenly appears in front of the vehicle, moving from the right side of the lane to the left, at a speed of 5km/h. The driver is not prompted to take over. Reference to these test scenario is illustrated in Figure 2.

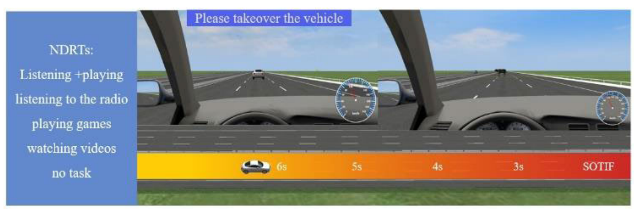

FIGURE 2. General layout of the test scenario.

This trial used a mixed design of 5 (5 NDRTs) x 5 (5 takeover scenarios) with 25 categories of trials. The NDRTs were the between-group variables and the takeover request prompts were the within-group variables. In the trials, each of the 5 takeover scenario trials completed while each test group performed one non-driving related task.

B. APPARATUS AND PARTICIPANTS

The study was conducted with a simple stationary driving simulator, as shown in Figure 3. The test scenario is displayed by the driving simulation display device. The driver controls the simulated vehicle via the Fanatec steering wheel-pedal set. The simulator was set to automatic mode, and when the driver developed a desire to take over, the red button next to the steering wheel could be pressed to switch between

FIGURE 3. Driving simulation equipment.

automatic and manual modes. The tests were carried out using UC-win/Road software for scenario building. The data was collected based on the Log plug-in provided by UC-win/Road and the driving data was collected at 25 Hz.

A total of 62 subjects completed the driving simulation test (18 females, 44 males). The mean age of the subjects was 35.29 years (SD = 6.88), the mean driving experience was 9.85 years (SD = 5.21) and the mean driving mileage was 38,700 km (SD = 15,500).

C. PROCEDURE

Prior to the formal trial participants were asked to sign an informed consent form and to complete a demographic questionnaire mainly including information on the participant's gender, age, driving experience, driving mileage and so on. Then, participants were trained in NDRTs and simulator driving, and were trained in driving simulations of other scenarios before the start of the formal trial. During the formal test, the automatic vehicle driving and NDRTs start simultaneously. The driver takes over according to the different takeover scenarios. Participants completed a subjective rating scale at the end of the trial.

This research was conducted in accordance with the guidelines of the "Helsinki Declaration" and was approved by the Review Committee of the School of Architecture and Transportation Engineering, Guilin University of Electronic Technology (Approval Number: 2021PS03). Informed consent was obtained from all participants involved in the study.

D. EVALUATION METRICS

We developed two types of features, takeover impact features and risk indicator features. The takeover impact features were extracted o provide a thorough and multi-perspective analysis of takeover performance, including six factors: driver gender, age, driving experience, driving mileage, takeover scenario, and NDRTs. Risk indicator features were used to directly characterize takeover performance, including takeover reaction time, steering reaction time, minimum TTC, and maximum synthetic acceleration. These metrics have also been widely used in previous studies [36]. Descriptions of each risk indicator feature are shown in Table 2.

TABLE 2. Risk indicator characteristics and descriptions.

Metrics	Definition
Driver Take-Over Performance	
take-over reaction time	Time elapsed between TOR and deactivation of the automation (by button press, brake pedal or steering).
steering reaction time	Time elapsed between TOR and turning the steering wheel to take over the vehicle.
vehicle take-over steady state	
minimum TTC	Minimum time-to-collisions to the broken-down vehicle in the avoiding scenarios.
maximum synthetic acceleration	Maximum arithmetic square root of the sum of the squared lateral and longitudinal accelerations before successful avoidance or accident after TOR.

III. MODEL DEVELOPMENT

A. FEATURE SUBJECTIVE WEIGHTING

The Analytic Network Process (ANP) is an extended development of the Analytic Hierarchy Process (AHP), proposed for analyzing multi-criteria decision-making in non-independent hierarchical structures. Unlike the AHP method, which decomposes decision problems hierarchically [37], [38], ANP considers the complex and interrelated relationships among decision elements. It also possesses the ability to simultaneously apply qualitative and quantitative attributes, enabling it to handle multiple, correlated, and conflicting indicators. ANP constructs comprehensive analytical models with a network structure, allowing for the clear reflection and quantification of inherent relationships among various indicators and factors. It is suitable for addressing comprehensive issues involving the integration of different-dimensional composite indicators for comprehensive evaluation.

Using the traditional Delphi method to construct the discriminant matrix in the ANP model may introduce significant subjective biases into indicator weights, resulting in subjective errors in the final weight outcomes. Therefore, subjective risk indicator characteristics are evaluated using the relative weight expert mean confidence method [39], which aims to reduce the impact of subjectivity in the calculations while reflecting the significance of the characteristics themselves [40]. The improved ANP model based on weighted means is shown in Figure 4.

Using the improved ANP method to determine the subjective weights of risk indicator characteristics, the calculation process consists of the following main steps: Employing a number of experts to make a two-by-two comparison of the network layer elements, build a standard judgement matrix of 1-9 scales under the control layer and network layer indicators, and finally the average evaluation value of the importance of the network layer elements obtained

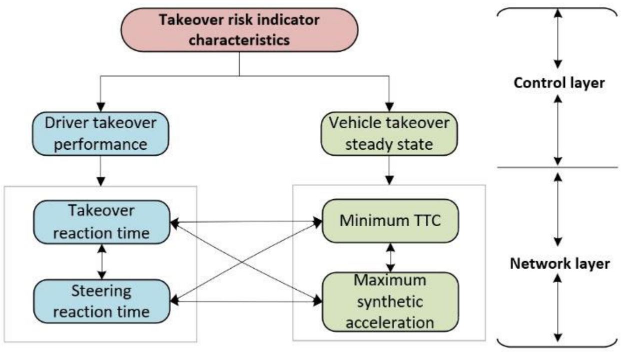

FIGURE 4. ANP evaluation model.

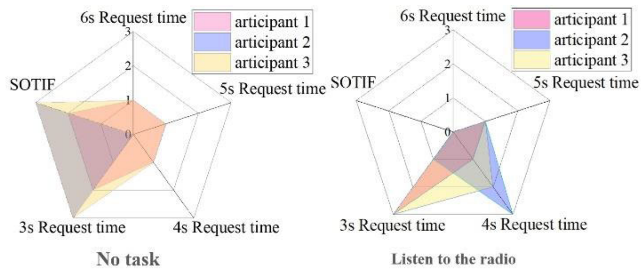

FIGURE 6. Results of the VAS assessment for selected drivers.

and tested for consistency. The eigenvectors of each matrix are calculated to build the unweighted supermatrix of the takeover risk assessment. The weights of the indicators calculated by the improved ANP method are shown in Table 3. using the improved ANP model to assign weights to the risk indicator data, reflecting the takeover data information while focusing on the importance of the risk indicator characteristics.

TABLE 3. Indicator weights based on the improved ANP method.

Elements	Driver takeover performance		Vehicle takeover steady state	
	reaction time	steering time	minimum TTC	maximum synthetic acceleration
v_1 Driver takeover performance	0.458850	0.412820	0.102635	0.025695
v_2 Vehicle takeover steady state	0.112483	0.030850	0.454348	0.402320
β_i Integration weight	0.285667	0.221835	0.278491	0.214007

B. QUANTIFYING TAKEOVER RISK

The dichotomous qualitative accident data were quantified using the Visual Analogue Scale/Score (VAS) pain scoring criteria. Reference to the VAS assessment methodology is elucidated in Figure 5. The subjective rating of accident rate was based on the driver's takeover process as the scoring basis, the importance of each characteristic obtained by the improved ANP method as the scoring criteria, and collision situation as the scoring constraint [41]. Some VAS quantitative data are shown in Figure 6.

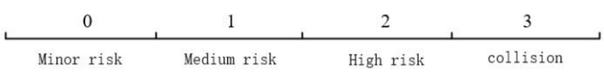

FIGURE 5. VAS assessment methodology.

C. XGBOOST ALGORITHM TRAINING

XG-Boost have been proved to have great prediction performance for classification problems [42]. The XG-Boost algorithm is a gradient boosting tree algorithm that supports parallel computation and is based on the Gradient Boosting Decision Tree (GBDT). For example, the second-order Taylor expansion loss function is used to improve the computational accuracy; the regularization term is introduced to simplify the model to prevent overfitting and improve the running speed; and the Blocks storage structure is used to achieve parallel computation.

Assuming a total of k trees, the model prediction $\hat{y}_i^{(t)}$ for round t can be expressed as (1).

$$\hat{y}_i^{(\mathrm{t})} = \sum\nolimits_{k=1}^{t} f_k\left(\mathrm{x_i}\right) = \hat{y}_i^{(\mathrm{t}-1)} + f_t\left(\mathrm{x_i}\right) \tag{1}$$

where: t is the number of iterations; $f_k(\mathrm{x_i})$ is the prediction of the k-th tree for variable $\mathrm{x_i}$; $\hat{y}_i^{(\mathrm{t}-1)}$ is the model prediction for round $\mathrm{t}-1$; and $f_t(\mathrm{x_i})$ is the tree function for round t. The objective function and the regularization term $\Omega(f_t)$ can be expressed as (2) and (3).

$$Obj = \sum\nolimits_{i=1}^{n} l\left(y_i, \hat{y}_i\right) + \sum\nolimits_{k=1}^{t} \Omega\left(f_k\right) \tag{2}$$

$$\Omega\left(f_t\right) = \gamma T + \frac{1}{2}\lambda \sum\nolimits_{j=1}^{T} w_j^2 \tag{3}$$

where: $l(y_i, \hat{y}_i)$ is the loss function; γ and λ are adjustment parameters to prevent over-fitting of the model; T is the number of leaf nodes; and w is the leaf node weight. A Taylor expansion is used for the objective function to increase the speed and accuracy of the gradient descent, and the expanded objective function is shown in (4).

$$Obj^t = \sum\nolimits_{j=1}^{T} [G_j w_j + \frac{1}{2}(H_j + \lambda) w_j^2] + \gamma T \tag{4}$$

where: $G_j = \sum_{i \epsilon I_j} g_i; H_j = \sum_{i \epsilon I_j} h_i$; g_i and h_i are the first- and second-order output loss gradients for the ith sample, respectively. Thus, the problem is transformed into the problem of solving the optimal value of (4).

IV. RESULTS AND DISCUSSION

A. HYPER-PARAMETER TUNING

A driver takeover risk assessment model was developed based on the XGBoost algorithm. Takeover influencing factors and weighted risk indicator data were used as input, and quantified takeover risk level as output, with an 8:2 training-to-test data ratio. To reduce model complexity, ensure high accuracy, and prevent overfitting, a grid search method grid search (GridSearch) was used to rank and combine the variables in the list of hyperparameter combinations to be tested, each combination was iterated through, and the model object with the highest mean score hyperparameter combination was selected as the best choice via 5-fold cross-validation. The results of model parameter selection are shown in Table 4.

TABLE 4. Optimal model parameters.

Parameters	Range	Optimum value
learning rate	[0, 1, 0.1]	0.5
n_estimators	[0, 500, 100]	100
min_child_weight	[1, 10, 1]	1
max_depth	[1, 50, 1]	10
Colsample bytree	[0.5, 1, 0.1]	0.9
gamma	[0, 1, 0.1]	0.5
subsample	[0, 1, 0.1]	0.7

B. FEATURE IMPORTANCE AND CORRELATION ANALYSIS

Using the weight index in the XGBoost algorithm, feature variable importance analysis is performed, generating a relative importance score for each feature based on the segmentation weights. These scores represent the number of times a feature is used in all trees and reflect the usefulness of the feature in constructing the boosted tree in XGBoost. The higher the score, the greater the importance of the feature to the final prediction result. Based on the obtained scores, the feature importance is screened, and the driver gender factor with the lowest score is removed. After this step, the model performance is significantly improved. Further reducing the number of features leads to a decline in classification performance. Therefore, nine feature variables as shown in Figure 7 are selected to construct the driver takeover risk assessment model. Among them, the risk indicator features with the greatest impact on the takeover risk level are the minimum TTC, maximum synthetic acceleration, takeover response time and steering reaction time, and the takeover influencing factors with the greatest impact on the takeover risk level are driving mileage, takeover scenario, driving experience, age and NDRTs.

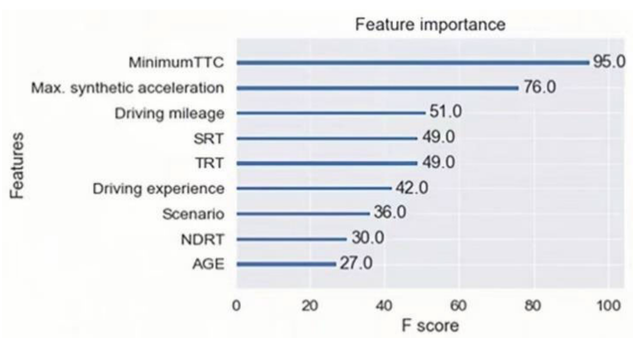

FIGURE 7. Feature importance.

The correlation between the characteristics affecting the level of takeover risk is visualized in Figure 8. Consistent with the feature importance results, the correlation between minimum TTC and takeover risk is the highest (r=−0.81), meaning that the level of takeover risk increases as minimum TTC decreases; takeover scenarios are negatively correlated with the minimum TTC and the degree of takeover risk, indicating that as the takeover prompt time continues to move back and until the SOTIF issue occurs, the minimum TTC decreases and increases the risk of takeover; takeover response time was more correlated with takeover risk than steering response time, both being negatively correlated with minimum TTC, slower takeover response producing smaller crash times; takeover response time and steering response time increased with higher task load for NDRTs and were negatively correlated with minimum TTC.

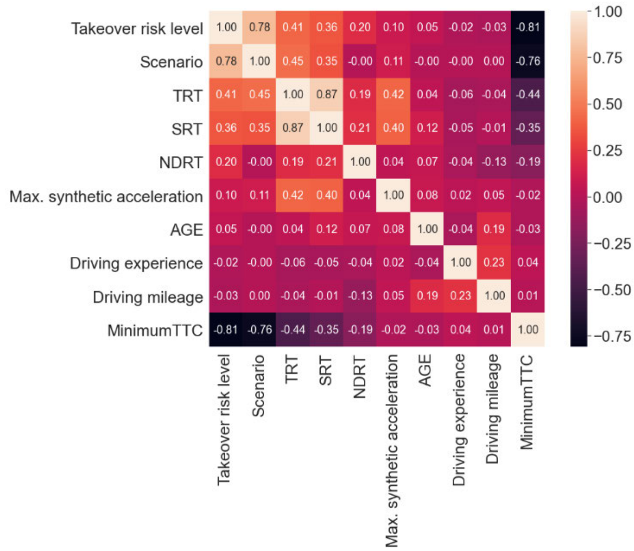

FIGURE 8. Feature importance.

C. MODEL PERFORMANCE COMPARISONS

In machine learning, the classification effectiveness of a model can be represented by four metrics. True Positive (TP): Refers to instances where the model correctly predicts positive samples as positive. True Negative (TN): Denotes instances where the model correctly predicts negative samples as negative. False Positive (FP): Represents instances where the model incorrectly predicts negative samples as positive. False Negative (FN): Signifies instances where the model incorrectly predicts positive samples as negative.

In this paper, accuracy, precision, recall, F1 score, and ROC curve are commonly used metrics for evaluating the performance and effectiveness of classification models. Higher values of the metrics represent higher performance of the model. They were formulated as follows.

Accuracy: Accuracy measures the proportion of correctly predicted samples to the total number of samples in a classification model. It is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Precision: Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive by a classification model. It is calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

Recall: Recall measures the proportion of correctly predicted positive instances out of all actual positive instances in a classification model. It is calculated as:

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

F1 Score: The F1 score is the harmonic mean of precision and recall, providing a balanced assessment of a model's accuracy and identification capability. It is calculated as:

$$F1Score = \frac{PrecisionRecall}{Precision + Recall} \tag{8}$$

ROC Curve: The ROC curve is a graphical representation of the true positive rate (TPR) and false positive rate (FPR) of a classification model at different thresholds. It is used to evaluate the model's performance across various thresholds, with the area under the curve (AUC) representing the model's overall performance.

Using the LightGBM, support vector machines (SVM) algorithm and XGboost model to compare the classification performance, the results of the model performance evaluation based on the accuracy, precision, recall and F1 values are shown in Table 5. The classification performance of the XGboost model was the best, with 87.1% accuracy, 83.4% precision, 83.5% recall and 83.4% F1 value. The accuracy rates for different takeover risk classes are shown in Figure 9. The accuracy rate of identifying low takeover risk level is 96.31%, the accuracy rate of identifying medium takeover risk level is 75.46%, the accuracy rate of identifying high takeover risk level is 95.95%, and the accuracy rate of identifying collisions is 100%.

TABLE 5. Model classification performance.

Models	Accuracy	Precesion	Recall	F1 value
XGboost	0.871	0.834	0.835	0.834
LightGBM	0.822	0.758	0.755	0.756
SVM	0.629	0.483	0.5553	0.502

D. LIMITATIONS AND FUTURE DIRECTIONS

This paper conducted driving simulation trials and constructed a takeover risk assessment model to evaluate takeover risk. The driving simulation platform technology offers advantages such as ease of operation, accurate data

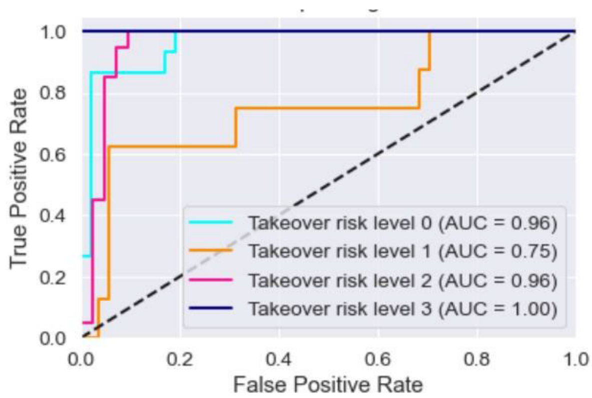

FIGURE 9. ROC curve.

extraction, and high safety. However, it differs from real vehicle tests in terms of authenticity. For instance, drivers may have poorer perception of vehicle speed in driving simulation scenarios, potentially leading to more urgent takeovers and increased takeover risk. Additionally, real driving environments are more variable than simulated scenes, and environmental factors may also influence takeover risk levels. Therefore, future research could investigate the levels and differences in takeover risk between simulated and actual driving under different driving tasks and takeover request prompts. Furthermore, the simulated experiments in this study primarily focused on the immediate impact of takeover prompts and non-driving related tasks on driver takeover risk. There is a lack of long-term research on takeover risk regarding driver cognitive decision-making and human-machine interaction adaptability. However, in traditional driving modes, the duration of driving may affect drivers' cognitive and physiological loads. Therefore, future studies should explore the impact of repeated human-machine interactions in the same scenario on risk levels, integrating investigations into drivers' psychological and physiological indicators related to cognitive resources and decision-making during takeover processes.

V. CONCLUSION

This study employed driving simulation techniques to design a hybrid experiment. Data on 310 takeover impact features and risk indicator features were collected from 62 drivers. A driver takeover risk assessment model based on the improved ANP method and XGBoost algorithm is proposed for predicting and analysing driver takeover performance. Our proposed method amalgamates subjective and objective evaluation advantages, accentuating the importance of risk indicator features while reflecting takeover data information. Moreover, our method exhibits superior accuracy. The main conclusions are presented as follows:

The highest correlation between minimum TTC and takeover risk level was found among the risk indicator features.

The highest correlation between scenario factors and takeover risk level and minimum TTC was found among the takeover influence features, which leads to the conclusion that scenario factors have the greatest influence on driver takeover safety.

NDRTs had some correlation with takeover reaction time, steering reaction time and minimum TTC, but the effect was minor.

The XGBoost algorithm-based takeover risk assessment model achieves a classification accuracy of 87.1%, an accuracy rate of 83.4%, a recall rate of 83.5%, and an F1 value of 83.4%. Outperforming LightGBM and SVM algorithms, it effectively discriminates driver takeover risk levels by utilizing risk indicator features and takeover impact feature data.

REFERENCES

[1] M. Delmas, V. Camps, and C. Lemercier, "Effects of environmental, vehicle and human factors on comfort in partially automated driving: A scenario-based study," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 86, pp. 392–401, Apr. 2022, doi: 10.1016/j.trf.2022.03.012.

[2] S. Ma and W. Zhang and J. Shi and Z. Yang, "The human factors of the take-over process in conditional automated driving based on cognitive mechanism," *Adv. Psychol. Sci.*, vol. 28, pp. 150–160, Jan. 2020, doi: 10.3724/SP.J.1042.2020.00150.

[3] H. Zhou, X. Li, X. He, P. Li, L. Xiao, and D. Zhang, "Research on safety of the intended functionality of automobile AEB perception system in typical dangerous scenarios of two-wheelers," *Accident Anal. Prevention*, vol. 173, Aug. 2022, Art. no. 106709, doi: 10.1016/j.aap.2022.106709.

[4] B. Wang, Y. Luo, Z. Zhong, and K. Li, "Risk reduction for safety of the intended functionality of CACC with complex uncertainties: A cooperative robust non-fragile fault tolerant strategy," *Transp. Res. C, Emerg. Technol.*, vol. 144, Nov. 2022, Art. no. 103885, doi: 10.1016/j.trc.2022.103885.

[5] L. Petersen, L. Robert, X. J. Yang, and D. Tilbury, "Situational awareness, driver's trust in automated driving systems and secondary task performance," *SAE Int. J. Connected Automated Vehicles*, vol. 2, no. 2, pp. 129–141, May 2019, doi: 10.4271/12-02-02-0009.

[6] J. J. So, S. Park, J. Kim, J. Park, and I. Yun, "Investigating the impacts of road traffic conditions and driver's characteristics on automated vehicle takeover time and quality using a driving simulator," *J. Adv. Transp.*, vol. 2021, pp. 1–13, Jun. 2021, doi: 10.1155/2021/8859553.

[7] W. Morales-Alvarez, O. Sipele, R. Léberon, H. H. Tadjine, and C. Olaverri-Monreal, "Automated driving: A literature review of the take over request in conditional automation," *Electronics*, vol. 9, no. 12, p. 2087, Dec. 2020, doi: 10.3390/electronics9122087.

[8] D. Choi, T. Sato, T. Ando, T. Abe, M. Akamatsu, and S. Kitazaki, "Effects of cognitive and visual loads on driving performance after take-over request (TOR) in automated driving," *Appl. Ergonom.*, vol. 85, May 2020, Art. no. 103074, doi: 10.1016/j.apergo.2020.103074.

[9] Q. Lin, S. Li, X. Ma, and G. Lu, "Understanding take-over performance of high crash risk drivers during conditionally automated driving," *Accident Anal. Prevention*, vol. 143, Aug. 2020, Art. no. 105543, doi: 10.1016/j.aap.2020.105543.

[10] C. Wu, H. Wu, N. Lyu, and M. Zheng, "Take-over performance and safety analysis under different scenarios and secondary tasks in conditionally automated driving," *IEEE Access*, vol. 7, pp. 136924–136933, 2019, doi: 10.1109/ACCESS.2019.2914864.

[11] E. Shi and K. Bengler, "Non-driving related tasks' effects on takeover and manual driving behavior in a real driving setting: A differentiation approach based on task switching and modality shifting," *Accident Anal. Prevention*, vol. 178, Dec. 2022, Art. no. 106844, doi: 10.1016/j.aap.2022.106844.

[12] J. W. Niu, X. M. Zhang, Y. P. Sun, and H. Qin, "Analysis of the driving behavior during the takeover of automatic driving vehicles in dangerous traffic situations," *China J. Highway Transp.*, vol. 31, no. 6, pp. 272–280, 2018.

[13] G. Q. Lu, P. Y. Zhao, Z. J. Wang, and Q. F. Lin, "The effect of visual subtasks on young drivers' takeover time in autonomous driving," *China J. Highway Transp.*, vol. 31, pp. 165–171, Jan. 2018.

[14] P. Bazilinskyy, S. M. Petermeijer, V. Petrovych, D. Dodou, and J. C. F. de Winter, "Take-over requests in highly automated driving: A crowdsourcing survey on auditory, vibrotactile, and visual displays," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 56, pp. 82–98, Jul. 2018, doi: 10.1016/j.trf.2018.04.001.

[15] X. Tan and Y. Zhang, "The effects of takeover request lead time on drivers' situation awareness for manually exiting from freeways: A web-based study on level 3 automated vehicles," *Accident Anal. Prevention*, vol. 168, Apr. 2022, Art. no. 106593, doi: 10.1016/j.aap.2022.106593.

[16] C. Xu, P. Li, Y. Li, N. Merat, Z. Lu, and X. Guo, "Takeover performance and workload under varying automation levels, time budget and road curvature," in *Proc. IEEE Asia–Pacific Conf. Image Process., Electron. Comput. (IPEC)*, Apr. 2022, pp. 1379–1385, doi: 10.1109/IPEC54454.2022.9777353.

[17] D. Shi, T. Wang, Y. Chen, and C. Liu, "Exploring the effects of request time, secondary task, and take-over mode on take-over performance," in *Proc. 6th Int. Conf. Intell. Transp. Eng.* (Lecture Notes in Electrical Engineering), vol. 901, Z. Zhang, Ed. Singapore: Springer, 2022, pp. 993–1006, doi: 10.1007/978-981-19-2259-6_88.

[18] S.-A. Kaye, S. Demmel, O. Oviedo-Trespalacios, W. Griffin, and I. Lewis, "Young drivers' takeover time in a conditional automated vehicle: The effects of hand-held mobile phone use and future intentions to use automated vehicles," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 78, pp. 16–29, Apr. 2021, doi: 10.1016/j.trf.2021.01.012.

[19] S. Li, P. Blythe, Y. Zhang, S. Edwards, W. Guo, Y. Ji, P. Goodman, G. Hill, and A. Namdeo, "Analysing the effect of gender on the human–machine interaction in level 3 automated vehicles," *Sci. Rep.*, vol. 12, no. 1, p. 11645, Jul. 2022, doi: 10.1038/s41598-022-16045-1.

[20] H. Loeb, A. Belwadi, J. Maheshwari, and S. Shaikh, "Age and gender differences in emergency takeover from automated to manual driving on simulator," *Traffic Injury Prevention*, vol. 20, pp. 163–165, Nov. 2019, doi: 10.6084/m9.figshare.10079171.v2.

[21] S. Li, P. Blythe, Y. Zhang, S. Edwards, J. Xing, W. Guo, Y. Ji, P. Goodman, and A. Namdeo, "Should older people be considered a homogeneous group when interacting with level 3 automated vehicles?" *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 78, pp. 446–465, Apr. 2021, doi: 10.1016/j.trf.2021.03.004.

[22] S. Li, P. Blythe, W. Guo, and A. Namdeo, "Investigation of older driver's takeover performance in highly automated vehicles in adverse weather conditions," *IET Intell. Transp. Syst.*, vol. 12, no. 9, pp. 1157–1165, Nov. 2018, doi: 10.1049/iet-its.2018.0104.

[23] S. Li, P. Blythe, W. Guo, A. Namdeo, S. Edwards, P. Goodman, and G. Hill, "Evaluation of the effects of age-friendly human-machine interfaces on the driver's takeover performance in highly automated vehicles," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 67, pp. 78–100, Nov. 2019, doi: 10.1016/j.trf.2019.10.009.

[24] Y. Wu, K. Kihara, Y. Takeda, T. Sato, M. Akamatsu, and S. Kitazaki, "Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving," *Accident Anal. Prevention*, vol. 124, pp. 202–209, Mar. 2019, doi: 10.1016/j.aap.2019.01.013.

[25] B. Zhang, J. de Winter, S. Varotto, R. Happee, and M. Martens, "Determinants of take-over time from automated driving: A meta-analysis of 129 studies," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 64, pp. 285–307, Jul. 2019, doi: 10.1016/j.trf.2019.04.020.

[26] F. Chen, G. Lu, Q. Lin, J. Zhai, and H. Tan, "Are novice drivers competent to take over control from level 3 automated vehicles? A comparative study with experienced drivers," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 81, pp. 65–81, Aug. 2021, doi: 10.1016/j.trf.2021.05.012.

[27] C. Gold, R. Happee, and K. Bengler, "Modeling take-over performance in level 3 conditionally automated vehicles," *Accident Anal. Prevention*, vol. 116, pp. 3–13, Jul. 2018, doi: 10.1016/j.aap.2017.11.009.

[28] R. Happee, C. Gold, J. Radlmayr, S. Hergeth, and K. Bengler, "Take-over performance in evasive manoeuvres," *Accident Anal. Prevention*, vol. 106, pp. 211–222, Sep. 2017, doi: 10.1016/j.aap.2017.04.017.

[29] K. Wiedemann, F. Naujoks, J. Wörle, R. Kenntner-Mabiala, Y. Kaussner, and A. Neukum, "Effect of different alcohol levels on take-over performance in conditionally automated driving," *Accident Anal. Prevention*, vol. 115, pp. 89–97, Jun. 2018, doi: 10.1016/j.aap.2018.03.001.

[30] H. Wu, C. Wu, N. Lyu, and J. Li, "Does a faster takeover necessarily mean it is better? A study on the influence of urgency and takeover-request lead time on takeover performance and safety," *Accident Anal. Prevention*, vol. 171, Jun. 2022, Art. no. 106647, doi: 10.1016/j.aap.2022.106647.

[31] H. C. Xian, L. S. Jin, K. Y. Li, and L. L. Gao, "A sub-task driving safety evaluation model based on network analysis method," *J. Chang'an Univ.*, vol. 35, pp. 124–128, Jan. 2015.
[32] Q. F. Lin, Z. J. Wang, and G. Q. Lu, "Takeover safety evaluation model for L3 level self-driving cars," *Automot. Eng.*, vol. 41, pp. 1258–1264, Jan. 2019.
[33] A. Ouadah, "Pipeline defects risk assessment using machine learning and analytical hierarchy process," in *Proc. Int. Conf. Appl. Smart Syst. (ICASS)*, Nov. 2018, pp. 1–6, doi: 10.1109/ICASS.2018.8651970.
[34] C. Gasne, L. Paire-Ficout, S. Bordel, S. Lafont, and M. Ranchet, "Takeover performance of older drivers in automated driving: A review," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 87, pp. 347–364, May 2022, doi: 10.1016/j.trf.2022.04.015.
[35] A. Eriksson, V. A. Banks, and N. A. Stanton, "Transition to manual: Comparing simulator with on-road control transitions," *Accident Anal. Prevention*, vol. 102, pp. 227–234, May 2017, doi: 10.1016/j.aap.2017.03.011.
[36] M. Jin, G. Lu, F. Chen, X. Shi, H. Tan, and J. Zhai, "Modeling takeover behavior in level 3 automated driving via a structural equation model: Considering the mediating role of trust," *Accident Anal. Prevention*, vol. 157, Jul. 2021, Art. no. 106156, doi: 10.1016/j.aap.2021.106156.
[37] A. Ouadah, A. Hadjali, F. Nader, and K. Benouaret, "SEFAP: An efficient approach for ranking skyline web services," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 2, pp. 709–725, Feb. 2019, doi: 10.1007/s12652-018-0721-7.
[38] Y. Sun and C. L. Giles, "Popularity weighted ranking for academic digital libraries," in *Advances in Information Retrieval* (Lecture Notes in Computer Science), G. Amati, C. Carpineto, and G. Romano, Eds. Berlin, Germany: Springer, 2007, pp. 605–612, doi: 10.1007/978-3-540-71496-5_57.
[39] J. Q. Chen, T. Tang, and L. Yuan, "Improved ANP model for evaluating the control capability of wireless occlusion centres," *J. Transp. Eng.*, vol. 15, pp. 108–118, Jan. 2015.
[40] M. Almulla, H. Yahyaoui, and K. Al-Matori, "A new fuzzy hybrid technique for ranking real world web services," *Knowl.-Based Syst.*, vol. 77, pp. 1–15, Mar. 2015, doi: 10.1016/j.knosys.2014.12.021.
[41] Z. Huang, I. V. Kohler, and F. Kämpfen, "A single-item visual analogue scale (VAS) measure for assessing depression among college students," *Community Mental Health J.*, vol. 56, no. 2, pp. 355–367, Feb. 2020, doi: 10.1007/s10597-019-00469-7.
[42] X. Shi, Y. D. Wong, M. Z.-F. Li, C. Palanisamy, and C. Chai, "A feature learning approach based on XGBoost for driving assessment and risk prediction," *Accident Anal. Prevention*, vol. 129, pp. 170–179, Aug. 2019, doi: 10.1016/j.aap.2019.05.005.

TAO WANG received the Ph.D. degree in traffic engineering from Southeast University, in 2017. He is currently the Dean of the School of Architecture and Traffic Engineering, Guilin University of Electronic Technology, and the Director of Guangxi Key Laboratory of Intelligent Transportation System. His research interests include urban traffic planning and management, intelligent networked transportation, and traffic behavior and safety.

YAXI HAN is currently pursuing the master's degree in transportation engineering with Guilin University of Electronic Technology. She was involved in several horizontal and vertical projects during her graduate studies. Her research interests include traffic planning and management, traffic behavior and safety, traffic information, and intelligent transportation systems.

WENYONG LI received the Ph.D. degree in transport engineering from Southeast University, in 2007. He is currently the President of Guangxi Transportation Vocational and Technical College and the Postgraduate Supervisor of Guilin University of Electronic Technology. His research interests include traffic planning and management, intelligent transport systems, traffic safety, and public transport.

XIAOFEI YE received the Ph.D. degree in transportation engineering from Southeast University, in 2013. He is currently an Assistant Professor with the Department of Logistics Management, Faculty of Maritime and Transportation, Ningbo University. His research interests include parking planning and management, nonmotorized vehicle and pedestrian traffic flow theory, traffic reliability, and advanced traffic information systems.

QUAN YUAN received the Ph.D. degree in vehicle engineering, in 1997. He is currently the Associate Director of the Institute of Intelligent Mobility, School of Vehicle and Mobility, Tsinghua University. His research interests include intelligent networked vehicles and smart shared mobility, traffic safety risks for vulnerable road users, autonomous driving risk assessment and collision-avoidance decisions, and V2I traffic accident reconstruction analysis.