

Study on a Risk Model and Target Detection for Prediction and Avoidance of Unmanned Environmental Hazard

Chengqun Qiu (✉ ujsqcq@126.com)

Yancheng Teachers University

Yuan Zhong

Yancheng Teachers University

Jie Ji

Chinese Academy of Sciences

Shuai Zhang

Yancheng Teachers University

Hui Zhang

Yancheng Teachers University

Shiqiang Zhao

Yancheng Teachers University

Mingyu Meng

Tokyo Institute of Technology

Research Article

Keywords: environmental hazard prediction, unmanned driving system, target detection, depth learning algorithm, SSD target detection method, method of combining the risk model

Posted Date: August 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-738182/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 Study on a risk model and target detection for prediction 2 and avoidance of unmanned environmental hazard

3 Chengqun Qiu^{1,2,*}, Yuan Zhong¹, Jie Ji³, Shuai Zhang¹, Hui Zhang¹, Shiqiang Zhao¹, Mingyu Meng⁴

4 ¹School of Clean Energy and Electrical Engineering, Yancheng Teachers University, Yancheng 224007, China

5 ²School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China

6 ³Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, 230031, China

7 ⁴Interdisciplinary Graduate School of Science & Engineering, Tokyo Institute of Technology, Yokohama 2268502, Japan

8 Corresponding author at: School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China

9 *E-mail address:* ujsqcq@126.com (C.Qiu)

10 Abstract

11 Comprehensive research is conducted on the design and control of the unmanned systems for
12 electric vehicles. The environmental risk prediction and avoidance system is divided into the
13 prediction part and the avoidance part. The prediction part is divided into environmental
14 perception, environmental risk assessment, and risk prediction. In the avoidance part, the
15 conservative driving strategy based on speed restriction is adopted according to the results of risk
16 prediction. Additionally, the core function is achieved through the target detection technology
17 based on deep learning algorithm and the data conclusion based on deep learning method.
18 Moreover, the location of bounding box is further optimized to improve the accuracy of SSD
19 target detection method based on solving the problem of unbalanced sample categories. Software
20 such as MATLAB and Carsim are applied in the system. From the comparison results of the
21 simulations of unmanned vehicles with or without a system, it that the system can provide
22 effective safety guarantee for unmanned driving.

23 **Keywords:** environmental hazard prediction; unmanned driving system; target detection; depth
24 learning algorithm; SSD target detection method; method of combining the risk model

25 Introduction

26 In the automobile industry, unmanned driving technology has attracted a great deal of attention in
27 recent years. It can fundamentally change the automobile industry and traffic systems. On the
28 other hand, it can also alleviate the problems of accidents, pollution, and congestion of existing
29 vehicles and the traffic ¹.

30 The commercialization of the unmanned driving should take safety as the premise and realize the
31 importance of the safe unmanned driving in the complex driving environment²⁻⁴, which is the
32 theme of this paper.

33 Anti-collision technology is one of the key points of unmanned research. Many achievements have
34 been made in the development of anti-collision technology, such as sensor information fusion,
35 anti-collision research, and anti-collision warning strategy⁴⁻⁶. However, it is still a long distance
36 from being completely practical considering the influence of multiple working conditions.

37 Some scholars have concluded the problems as follows:

38 Limited information fusion. At present, the research on sensor fusion is only the fusion between
39 two or three kinds of sensors, and the information of fusion cannot cover the working conditions
40 overall⁴⁻⁷. It is necessary to carry out the information fusion of various sensors and other data
41 sources in order to adapt to the actual driving conditions.

42 Multiple road conditions study is incomplete. No overall consideration is given to factors such as
43 road environment, weather conditions, the influence of personnel in the environment, and the
44 fastest response speed of vehicles⁵⁻⁸.

45 The early warning strategy needs improving. The present study basically takes distance as the
46 evaluation index. However, for the actual traffic situation, the process from safety to danger is a
47 gradual change, and multiple evaluation indexes should be used⁷⁻⁹.

48 To solve the problems above, this paper adopts the idea of dynamic risk assessment based on the
49 historical data of the environment and predicts the risk by priority based on the results of
50 environmental risk assessment¹⁰⁻¹². The integration of the booming internet big data industry and
51 electronic information engineering technology makes the risk assessment of traffic environment
52 no longer rely on manual rule setting and machine vision recognition, but can use the data from
53 navigation app and traffic department to realize the joint modeling and statistical analysis¹⁰.
54 Moreover, it is possible to dynamically assess the risk of the environment based on historical
55 circumstances and reapply the assessment results to the risk prediction of specific objectives in the
56 environment¹³⁻¹⁵. Therefore, this train of thought has high practical significance and application
57 value.

58 Target detection is the leading technology of hazard prediction. The current target detection is
59 mainly aimed at pedestrian, traffic sign, or obstacle¹⁴⁻¹⁶. In 2019, It proposed an improved
60 SSD_ARC algorithm for key target detection tasks in driving scenarios¹⁵⁻¹⁷. This method can
61 realize fast multi-objective recognition, semantic annotation, and positioning box selection.
62 Although it provides a general recognition framework, it does not involve the risk of identifying
63 the environment itself.

64 By contrast, this system makes up for this omission by adopting the idea of the priority of the big
65 data risk conclusion model and supplement of target detection, which has high practical
66 significance and application value¹⁸⁻²⁰. On this basis, this paper proposes an improved SSD
67 method by two steps:

68 First, the positive sample and negative sample imbalance of SSD are improved by FL loss
69 function.

70 The second is to improve the common boundary box selection and matching in the target detection
71 algorithm.

72 This paper presents an environmental hazard prediction and avoidance system. The system is
73 divided into two parts: the first part is the prediction part, which is divided into three levels,
74 including environmental perception, environmental risk model, and target detection. The second
75 part is the avoidance part. According to the results of hazard prediction, conservative driving
76 strategy based on speed limit is adopted. By using this system, vehicles can slow down in
77 high-risk areas or traffic complex environments and increase their speed when the risk is low.

78 The core of safe driving lies in avoiding danger. However, avoiding danger will inevitably affect
79 driving speed and comfort, especially avoiding environmental danger, which is mainly
80 accomplished by carrying out a defensive driving strategy. Therefore, the most important part of
81 this paper is the prediction of whether to use the defensive driving strategy. The prediction section
82 first identifies the environment, such as identifying the intersection, lane, parking lot, and
83 pedestrian crossing near the primary and secondary school campus, and then evaluates the risk and
84 gives the forecast target and priority according to the historical data. Finally, the risk index of the
85 target is predicted separately and evaluated synthetically.

86 At present, the visual algorithm of environment perception can complete the task of environment
87 recognition²¹⁻²⁵. By combining LBS positioning and other methods, the environmental information
88 can be preset in advance, and the recognition speed and accuracy of the visual algorithm can be
89 improved at the same time. Environmental risk assessment algorithm uses deep learning
90 technology, with the help of Internet open traffic accident database, the comprehensive analysis of
91 traffic accident affected factors in order to rank the dangerous objectives in the environment. It is
92 feasible to practice the idea, but predicting priorities will take a lot of testing to finalize. Besides,
93 under unknown circumstances, the hazard prediction algorithm has been realized in the conflicts
94 of people and cars. The risk prediction is still being explored and the risk error of traffic
95 environment under different time, weather, and other factors need further correcting. The thesis' s
96 structure is divided into four aspects:

97 Section1: This paper briefly introduces the development of automobile safety and anti-collision
98 technology and explains the importance of anti-collision technology to the driverless. Although
99 the environmental hazard prediction and avoidance system has not been developed, the
100 significance and prospect of this system are expounded.

101 Section2: There are four kinds of target detection methods commonly used in unmanned driving.
102 In this design, we will focus on the risk prediction based on deep learning.

103 Section3: Hardware and software design of environmental hazard prediction and avoidance
104 system.

105 Section4: The main function of the environmental hazard prediction avoidance system is to avoid
106 the risk. With MATLAB, CarSim software for simulation, we can eventually obtain the
107 experimental results to prove the feasibility of the system design implementation.

108

Nomenclature

Subscripts

C	real license plate size	C'	size of the object in the picture
d	euclidean geometric distance	fps	frame rate
P	frame storage coordinates	P'	frame storage coordinates
M_1	camera internal parameter matrix	M_2	camera position matrix
P_y	target of P on the Y-axis	Q	istance from the camera to the nearest point below
H	camera head height	V	pixel height of the image
h	camera height	v	pixel height coordinate of the target in the image
v_0	internal parameters provided for calibration	c	category confidence prediction
f_y		g	position parameter
σ	estimated uncertainty	x_g	real boundary box position on the ground
L	the predicted value of the position	f	camera focal length
x_e	estimated boundary box position and standard deviation	D	Kullback-Leibler divergence
σ_i	adjustable parameter for variable voting	L_{cls}	classification loss
L_{reg}	bounding box regression	t_{xi}^*, t_{yi}^*	deviation between the real and the anchor boundary frame
t_{xi}, t_{yi}	deviation between the prediction and the anchor boundary frame	x_i^*	true box position
x_i	bounding box position	y_i^*	
y_i			

Abbreviations

FL	Focal Loss	CNN	Convolutional Neural Network
RNN	Recurrent Neural Network	$LSTM$	Long Short-Term Memory
SSD	Single Shot MultiBox Detector	LBS	Location Based Service
BDS	BeiDou Navigation Satellite System	KL	Kullback-Leibler

109 **System model**

110 The realization of the environment awareness system mainly includes four steps: Firstly, it can
111 input the positioning data to the system with BDS/GPS satellite positioning, LBS positioning of
112 WIFI and base stations. Secondly, it can also use the electronic compass module to achieve
113 position refinement. Moreover, the environment prejudgment's realizing is based on position and
114 machine vision. Finally, the environmental data is output. Among them, the methods of satellite
115 positioning and electronic compass positioning are quite mature, but how to achieve
116 environmental judgment and corresponding risk assessment on this basis is the key problem to be
117 solved by the environmental awareness system. In this paper, a risk model is established based on
118 location, accident data, and a target detection algorithm through depth learning, which is proposed
119 to realize environmental judgment.

120 *Risk Model Based on Location and Accident Data*

121 According to the location information provided by the satellite and the electronic compass, it is
122 possible to make judgments on the types of nearby environment. There are six categories of
123 judgments: residential land, industrial land, public facilities land, commercial building land,
124 transportation facilities land, and road land.

125 On this basis, since most driverless vehicles using this system are running, making detailed
126 perception based on machine vision more necessary, which can be divided into two types:
127 intersection and road. Driving environment types are shown in Table 1. Intersections can be
128 divided into three types: three branches, four branches and, multiple branches. And roads can be
129 divided into four types: expressways, main roads, secondary roads, and branch roads. Allowing for
130 two points above, so as to obtain the judgment of the driving environment.

Classification of driving environment	Type of driving environment
Intersections	Three branches
Intersections	Four branches
Intersections	Multiple branches
Roads	Expressway
Roads	Main road
Roads	Secondary trunk road
Roads	Branches

131 **Table 1.** Driving environment types.

132 Based on the environment-aware data and the type of the nearby environment, the specific name
133 of the nearby environment can be obtained. At the same time, the system can carry out Internet

134 communication and obtain real-time traffic and weather conditions. Allowing for the above
 135 information, environmental risk can be judged from the following three aspects. First of all, risk
 136 judgment is ultimately to judge risk types and risk objectives. Secondly, they can be summarized
 137 as car-car conflict risk, car-person conflict risk, car-object conflict risk, and vehicle control risk.
 138 Eventually, risk objectives are visual objectives such as vehicles, pedestrians, bicycles and electric
 139 vehicles. Moreover, risk itself is divided into real risk and hidden risk, and real risk is the
 140 possibility of collision between risk target and vehicle on site. The hidden risk is difficult to
 141 confirm due to various reasons. But there is still the possibility of collision.

142 Location based

143 Risks based on nearby environment types are shown in Table 2. The types of nearby environment
 144 can be divided into several major categories such as residential, industrial, public facilities,
 145 commercial and transportation facilities environment. And the traffic facilities environment refers
 146 to bus stations, railway stations, airports, subway stations and other passenger transport hubs.

Type of nearby environment	Main risk types	Major risk objectives
Reside	Car-person conflict	Non-motor vehicle、 Pedestrians
Industry	Car-car conflict / Vehicle control	Vehicle
Public facilities	Car-person conflict	Non-motor vehicle、 Pedestrians
Commerce	Car-person conflict	Non-motor vehicle、 Pedestrians
Means of transportation	Car-car conflict / Car-person conflict	Vehicle、 Non-motor vehicle

147 **Table 2.** Risks based on nearby environment types.

148 Time based

149 Through the analysis of traffic data on roads, we can sort out the traffic flow of roads at different
 150 times. Generally speaking, large traffic volume and complicated traffic environment in each
 151 environment represent greater risks, such as working days, holidays and rush hours, could affect
 152 the traffic flow. For example, whether there is a road design for non-motor vehicle isolation design
 153 and the type of environment in which it is located all affect the complexity of traffic.

154 For working days, the risk of car-car conflict is more significant in most environments, and the
 155 same environment type is different in specific environment. For public facilities, full-time
 156 educational facilities have a significant risk of collision between people and vehicles after school.
 157 Besides, cultural facilities such as public libraries and museums have a higher risk during holidays,
 158 and medical facilities have different situations. For the business environment, different
 159 commercial districts have different traffic time distributions.

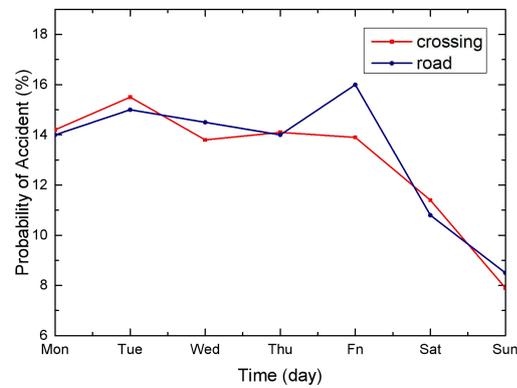
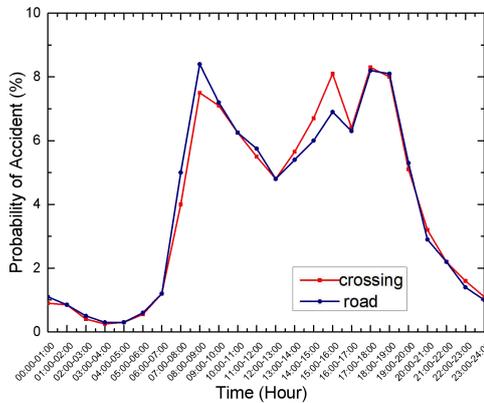
160 Therefore, this part of the program needs to input the current time. The specific risk type and
 161 priority of risk objectives will be determined through database query.

162 Based on the scene

163 Due to possible errors in location and database, the system confirms and supplements the on-site
 164 target detection. First, the targets, such as pedestrians and vehicles, are detected ahead. Compared
 165 with the above judgment results, the existing results are marked, or unexpected obstacles appear
 166 on the scene. This part is identified through machine vision to add the non-existing results and
 167 avoid the omission of risk targets.

168 Because the risk target is divided into real risk and possible risk, the existing risk target, such as
 169 vehicle, pedestrian which can be identified by the completed target detection algorithm, and risk
 170 weighting under space-time conditions is evaluated based on historical information.

171 Figure 1 and Figure 2 show the accident rate in hours and a week respectively based on the data
 172 from Shanghai. The appropriate model is generated with specific data.



173

174 **Figure 1.** Accident incidence rate in hours.

Figure 2. Accident incidence rate in a week.

175 The accident rate varies greatly with time, and vice versa. Weighted value table based on accident
 176 ratio is shown in Table 3. With the average value of 1 for each type, the accident data of the whole
 177 year are used for statistics, and the weighted value of each time period is re-weighted.

Day	Time								
	Type	01:00	02:00	08:00	09:00	10:00	11:00	12:00	18:00
Monday	Road	0.22	0.20	1.11	2.05	1.91	1.74	1.47	2.30
	Crossing	0.21	0.19	1.32	2.26	1.91	1.69	1.52	2.17
Tuesday	Road	0.24	0.22	1.21	2.24	2.09	1.90	1.60	2.51
	Crossing	0.28	0.21	1.42	2.42	2.04	1.81	1.63	2.33
Wednesday	Road	0.22	0.20	1.09	2.02	1.89	1.72	1.45	2.27
	Crossing	0.27	0.20	1.36	2.33	1.96	1.74	1.57	2.23
Thursday	Road	0.22	0.21	1.11	2.06	1.93	1.76	1.48	2.31

	Crossing	0.27	0.20	1.34	2.29	1.93	1.71	1.54	2.20
Friday	Road	0.22	0.20	1.08	2.01	1.87	1.71	1.44	2.25
	Crossing	0.29	0.22	1.46	2.50	2.11	1.87	1.68	2.40
Saturday	Road	0.17	0.16	0.87	1.60	1.50	1.37	1.15	1.80
	Crossing	0.20	0.14	0.99	1.68	1.42	1.26	1.13	1.62
Sunday	Road	0.13	0.12	0.63	1.16	1.09	0.99	0.83	1.31
	Crossing	0.16	0.12	0.81	1.38	1.17	1.03	0.93	1.33

178 **Table 3.** Weighted value table based on accident ratio.

179 Risks based on accident types are shown in Table 4. From the analysis of accident data, we can
180 know the accident type, accident vehicle, weather, time period and other information of the
181 accident according to the location. The accident type is the main sequence, which can be divided
182 into rear-end, reverse, reverse, door switch, traffic signal violation, non-yielding, other accidents
183 and other types. The correlation between each accident type and risk target and risk type can be
184 sorted. Also, the cause and party of the accident under this environment can be known at the same
185 time.

Accident type	Main risk types	Main causes of accidents
Rear-end	Car-car conflict/ Vehicle control	In front of the car brake/ground skid/visual blind area
Retrograde	Car-car / Car-person conflict	A blind spot/sudden rush
Astern	Car-car / Car-object conflict	A blind spot/sudden rush
Violation of traffic signal	Car-car / Car-person conflict	The signal lamp is shielded/blind/the other party suddenly rushes out.
Not yielding	Car-car / Car-person conflict	Too fast

186 **Table 4.** Risks based on accident types.

187 According to the current location, weather and time matching accidents, the priority of high-risk
188 accident objects is increased. According to the causes of accidents in the current environment, the
189 hidden risks are added and sorted. Therefore, the final program will output the risk target list data
190 with priority and hidden risks.

191 *Target detection method based on depth learning*

192 Deep learning trained models are applied to identify and detect the sequence of captured images.
193 And the algorithm is used to calculate the direction and speed of the target and the distance to
194 provide data for the next step.

195 The velocity prediction is realized by moving the European distance of the target center point
196 between adjacent Dayton. In short, there is a correspondence between the speed of the real world

197 image. If the target speed in the real world is fast, the speed in the adjacent pictures will be the
 198 same, so the speed can be obtained by finding the corresponding relation between the real speed
 199 and the video image speed. According to the shooting time of adjacent images, the frame rate, the
 200 moving distance of the target center and the moving speed in the images can be calculated.
 201 Because speed is affected by distance and time, but time is the same for real world and images, the
 202 converted distance is the most critical. What's more, the conversion relation can be obtained by
 203 using the real size and the image size. For unmanned video images, the license plate can be
 204 selected for objects of general size. With the help of the license plate width and the actual width in
 205 the image, the conversion ratio is obtained, thus obtaining the real distance and the real speed, and
 206 the relative velocity estimation formula of the target is as follows.

207 Ratio of image to real world:

$$208 \quad \text{Scale} = \frac{C}{C'} \quad (1)$$

209 Actual speed size:

$$210 \quad \text{Speed} = d \times \text{fps} \times \text{scale} \quad (2)$$

211 Where C is the real license plate size, C' is the size of the object in the picture, d is the euclidean
 212 geometric distance of the target moving in the image determined by the displacement of the center
 213 point, and fps is the frame rate. Since the velocity is vector, the velocity direction of the target
 214 should be obtained in addition to scalar. Firstly, image sequence groups within a period of time
 215 should be screened out. Secondly, the object center of the same target should be locked, and the
 216 moving direction of the object center in the sequence group should be determined to obtain the
 217 direction of instantaneous velocity.

218 For distance calculation, vision distortion and other issues should be considered first when CMOS
 219 sensors are used. The correction matrix and camera internal parameters could be obtained by using
 220 Matlab camera calibration toolbox and calibration function of OpenCV library. The details will
 221 not be described here due to the length of the paper.

222 The system applies a fixed device to perform a single visual distance algorithm. Through
 223 conversions from real-world to camera coordinate, camera to image coordinate conversion and
 224 image to frame storage coordinate, the conversion from real world to frame storage coordinate is
 225 realized:

$$226 \quad z_v \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & u_0 \\ 0 & s_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3)$$

227 where, (X,Y,Z) is the real world coordinate system, (X_v,Y_v,Z_v) is the camera coordinate
 228 system, (x_p,y_p) is the image coordinate, (s_x,s_y) is the unit of dividing pixels by millimeters, the
 229 origin of the fixed frame storage coordinate system is (u_0,v_0) , set any position as (u,v) , R is the

230 3x3 rotation matrix, T is the 3x1 translation matrix, and f is the camera focal length. Simplification
 231 can be done again:

$$232 \quad z_v \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4)$$

233 Finally reduced to:

$$234 \quad Z_v \times P = M_1 \cdot M_2 \cdot P' \quad (5)$$

235 P is the frame storage coordinates, P' is the real world coordinates, M_1 is the camera internal
 236 parameter matrix, M_2 is a camera position matrix.

237 Take the real situation as a profile of the Y-axis, set the P as the target, and project P_y on the
 238 Y-axis. After deduction, the distance formula can be obtained:

$$239 \quad d = h / \tan \left(\arctan \frac{Q}{h} + \arctan \frac{H}{2f} + \arctan \frac{P_y - y_0}{f} \right) \quad (6)$$

240 Take Q as the distance from the camera to the nearest point below, h as the camera height, H as the
 241 camera head height, (x_0, y_0) is the midpoint coordinate of the image. Using coordinate system
 242 conversion:

$$243 \quad d = h / \tan \left(\arctan \frac{Q}{h} + \arctan \frac{v}{2f_y} + \arctan \frac{v - v_0}{f_y} \right) \quad (7)$$

244 This time, V is the pixel height of the image, v is the pixel height coordinate of the target in the
 245 image, and v_0 and f_y are internal parameters provided for calibration.

246 At the same time, according to its own speed calculation, some risk targets have been or will be on
 247 the collision path, and this kind of realistic risk targets are marked as the highest priority. In
 248 addition, the priority is arranged in turn according to the speed and distance of the target.

249 System construction

250 The prediction part firstly recognizes and perceives the environment, such as identifying
 251 intersections, lanes, parking lots, crosswalks, the vicinity of primary and secondary schools, etc.,
 252 which is a risk model based on location and accident data. Secondly, the risk is evaluated
 253 according to historical data, that is, the risk model is used to give the prediction target and risk
 254 based on location and accident data. Finally, the target detection method based on depth learning is
 255 intended to detect the target and evaluate the risk index of the target. In a word, the system needs
 256 to solve the problems of "what is the current environment", "is there any risk in the environment",
 257 "what kind of risk is there", "the degree of danger of various risks" and "how to avoid it".

258 The trajectory of the risk target is predicted and tracked, and the braking distance is taken as the
 259 safety range for estimation. For hidden risks, the risk of ground skid caused by weather will

260 increase the braking distance, while the risk of line-of-sight problem assumes that objects with the
261 same speed as the vehicle are in the center of the shielding range, and estimates the safety index.

262 The parameters affecting the hazard value include the vehicle speed, braking performance, wet
263 skid degree of the road surface and the direction of the risk target speed. Therefore, the hazard
264 value should be obtained through comprehensive consideration of these parameters. According to
265 relevant documents, when emergency braking is used to avoid collision, deceleration greater than
266 5m/s^2 can be considered dangerous, 2 to 5m/s^2 is critical danger, and below 2m/s^2 it can be
267 considered safe. However, the road conditions will lead to a decrease in braking performance,
268 which is reflected in the deceleration at the maximum braking effect, referred to as the maximum
269 deceleration. Besides, the braking deceleration for any object should be less than the maximum,
270 especially for objects already in the field of view. It should be considered as appropriate even for
271 predicted objects that do not appear in the field of view. Therefore, the critical dangerous
272 deceleration should also give priority to the environmental ground friction coefficient. Figure 3
273 demonstrates the internal process shown in the flow chart of circumvention algorithm.



274

275 **Figure 3.** Flow chart of evasion algorithm.

276 **Algorithm model construction and demonstration based on deep** 277 **learning**

278 *Algorithm construction of environmental perception and hazard prediction*

279 The convolution neural network (CNN) and recurrent neural network (RNN) are used to complete
280 the task of environmental perception. The model based on LSTM variable gives different weights
281 to different features. It not only adapts to complex background, but also can deal with multiple
282 targets. What's more, a complete description can be obtained by using the end-to-end method on
283 the expression model, proposed by the Northwest University of Technology team in 2018.

284 Hazard prediction is divided into two parts: Target Detection and Hazard Degree prediction, in
285 which target detection is the application scenario of deep learning. Compared with the traditional
286 algorithm, algorithm based on deep learning has obvious advantages in detection accuracy and
287 efficiency. In this paper, an improved SSD-based target detection algorithm is proposed.

288 Extracting the feature information of important objects in traffic scene is the beginning of the
289 work. On the basis of supervised learning method, the attribute set is trained by multi-label
290 classification, and the attribute prediction is carried out by training the depth convolutional neural
291 network corresponding to the loss function.

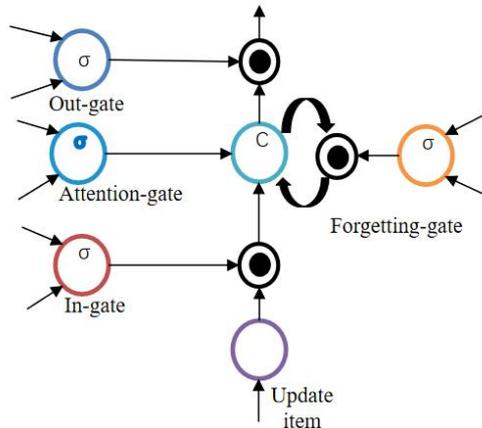
292 The supplementary description of environment perception belongs to the category of image
293 semantic recognition, and the method adopted belongs to the ‘end-to-end’.

294 The work of feature extraction is completed by CNN classification model. After classification, it is
295 represented by LSTM, an RNN variant model. It is particularly important to note that the LSTM
296 model is submitted not only to the extracted image features, but also to relevant information such
297 as color, focus range of attention, location, etc. The feature of this method lies in dividing attention
298 by color and weighting attention regions appropriately. However, the so-called color attention
299 weighting is to detect areas with relatively concentrated areas or large color changes of the same
300 color in the image, especially for red and colors with significant contrast. By the way, the
301 detection is realized by RGB color coding.

302 *The description of model*

303 LSTM is a special form of RNN network, whose structure has a storage unit for storing some
304 events with certain intervals and delays in the training process. The storage unit shown in Figure 4
305 regularly makes a trade-off of the content, and the trade-off is controlled by four gates. A
306 feature-based weight unit is generated in the gate control phase. Besides, the hidden layer state of
307 the previous node and the image features extracted by CNN are input to the unit, and the
308 stimulation features are analyzed by machine vision.

309 During the encoding phase, pictures and labels exist as vectors in the hidden layer state. Each
310 image extracts features with a trained VGG16 model. At the same time, the label vector is input
311 into the LSTM model through matrix transformation. During the decoding phase, the maximum
312 probability is obtained by multiplying the feature layer of the last layer of the hidden layer by the
313 seventh layer of the fully connected layer. After the comparison, the output model considers the
314 description to be the best match.

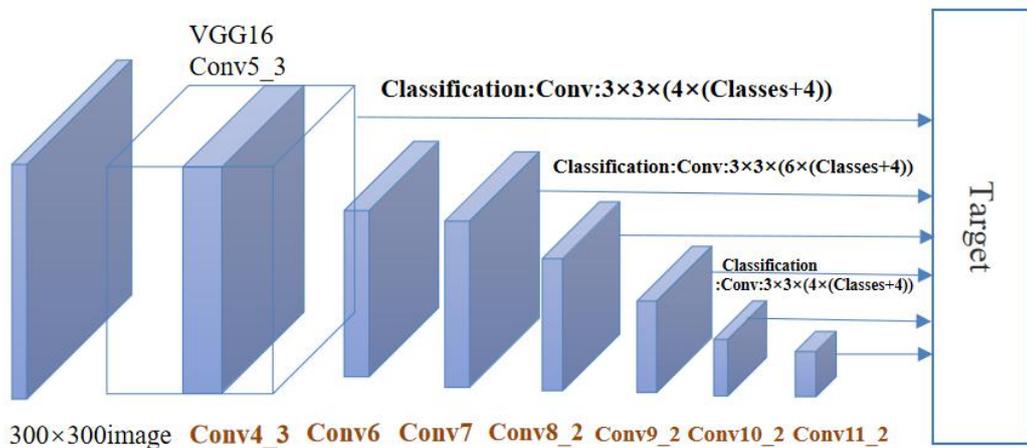


315

316 **Figure 4.** LSTM chart.

317 *Theoretical framework of SSD*

318 In the original SSD paper, the following structure is presented. SSD uses the feature pyramid
 319 structure for detection, which uses the characteristic smaps of conv4-3, 6-2, 7, 7-2, 8-2, 9-2. At the
 320 same time, position regression and softmax classification are performed. Figure 5 demonstrates
 321 that the SSDs can use VGG-16 as the base network. The feature extraction layer in the latter part
 322 is also predicted. In addition, the detection is performed not only on additional feature maps, but
 323 also on the underlying conv4-3 and 7-feature ditutations to achieve compatibility with small
 324 goals.



325

326 **Figure 5.** Schematic diagram of SSD architecture.

327 There are three core design concepts of SSD, as follows:

- 328 (a) There are two kinds of feature maps with multi-scale feature map: large feature map
 329 corresponds to small target and large target is responsible for small feature map.
- 330 (b) The feature map is extracted directly by convolution so that a large feature graph can be
 331 obtained with a relatively small convolution kernel.
- 332 (c) Setting a priori box each cell generates a priori box with different size, length and width. The

333 priori box serves as the baseline of the bounding box, and multiple priori frames are generated in
334 different ways during the training process.

335 Using VGG16 as the basic model, the SSD transforms the fully-connected layer into $3 \times 33 \times 3$
336 convolution layer CONV6 and $1 \times 11 \times 1$ convolution layer CONV7, and pool5 from 2×2 to 3×3 .
337 Then the FC8 and dropout layers are replaced by a series of convolution layers, and fine-tuned
338 them using the detection set. The Conv4 layer with a size of 38×38 in VGG16 will serve as the
339 first feature map for detection. But the layer data is too large to be normalized instead.

340 Five feature graphs were extracted from the new layers, namely Conv7、 , Conv8_2、 , Conv9_2,
341 Conv10_2 and Conv11_2, and the original layer of CONV4 was added, forming a total of six.
342 Their sizes are (38, 38), (19, 19), , (10, 10), (5, 5), (3, 3), (1, 1), (38, 38), (19, 19), (10, 10), (5, 5),
343 (3, 3), (1, 1). They have different priors, including size, length and width. What's more, as the size
344 of the feature map increases, the prior box size decreases:

345
$$w_k^a = s_k \sqrt{a_r}, h_k^a = s_k / \sqrt{a_r} \quad (8)$$

346 Then the results are obtained by convoluting the feature graph: Category Confidence and
347 bounding box position, each using a $3 \times 33 \times 3$ convolution to complete, the essence of SSD is
348 dense sampling.

349 *Algorithm training and improvement*

350 *Training*

351 Prior box matching

352 Before the work, the prior frame with the target or part of the target is retrieved, and the matching
353 boundary frame will enter the prediction phase. The first step of prior frame matching is to
354 confirm the largest prior frame for at least one frame to be identified. If it has a corresponding
355 target, it becomes a positive sample, otherwise it will be a negative sample. Secondly, if there is a
356 target matching degree greater than the threshold (generally 0.5) for the remaining negative
357 sample, the sample will become a positive sample. Moreover, targets may have multiple prior
358 frames that are not necessarily perfectly matched, but one prior frame cannot correspond to
359 multiple targets.

360 Loss function

361 The loss function can be understood as the weighted sum of confidence and position error:

362
$$L(x,c,l,g) = \frac{1}{N} [L_{conf}(x,c) + \alpha L_{loc}(x,l,g)] \quad (9)$$

363 where N represents the number of positive samples, $x_{ij}^p \in \{1,0\}$ is used as an indication parameter,
 364 and when $x_{ij}^p = 1$ is used, the i -th prior box matches the j -th target with category p . c represents the
 365 category confidence prediction. And L is the predicted value of the position, which is the position
 366 of the boundary of the target selected by the prior frame, and g represents its position parameter.
 367 The position error in the loss function only considers positive samples, which is defined by
 368 Smooth L1 loss as follows:

369
$$L(x,l,g) = \sum_{i \in Pos} \sum_{m \in \{cx,cy,wh\}} x_{ij}^k smooth_{L1} (l_i^m - \hat{g}_j^m) \quad (10)$$

370
$$\hat{g}_j^{cx} = \frac{(g_j^{cx} - d_i^{cx})}{d_i^w} \quad (11)$$

371
$$\hat{g}_j^{cy} = \frac{(g_j^{cy} - d_i^{cy})}{d_i^h} \quad (12)$$

372
$$\hat{g}_j^w = \log \frac{(g_j^w)}{d_i^w} \quad (13)$$

373
$$\hat{g}_j^h = \log \frac{(g_j^h)}{d_i^h} \quad (14)$$

374
$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (15)$$

375 The parameters are as follows:

$$376 \quad \widehat{g}_j^{cx} = \frac{(g_j^{cx} - d_i^{cx})}{\text{variance}} \quad (16)$$

$$377 \quad \widehat{g}_j^{cy} = \frac{(g_j^{cy} - d_i^{cy})}{\text{variance}} \quad (17)$$

$$378 \quad \widehat{g}_j^w = \frac{\log\left(\frac{g_j^w}{d_i^w}\right)}{\text{variance}} \quad (18)$$

$$379 \quad \widehat{g}_j^h = \frac{\log\left(\frac{g_j^h}{d_i^h}\right)}{\text{variance}} \quad (19)$$

380

381 For confidence error, it adopts softmax loss:

$$382 \quad L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\widehat{c}_i^p) - \sum_{i \in Neg} \frac{\Sigma(\widehat{c}_i^0) \text{ where } \frac{\exp(c_i^p)}{\Sigma_p(c_i^p) \exp}}{\Sigma_p(c_i^p) \exp} \log \quad (20)$$

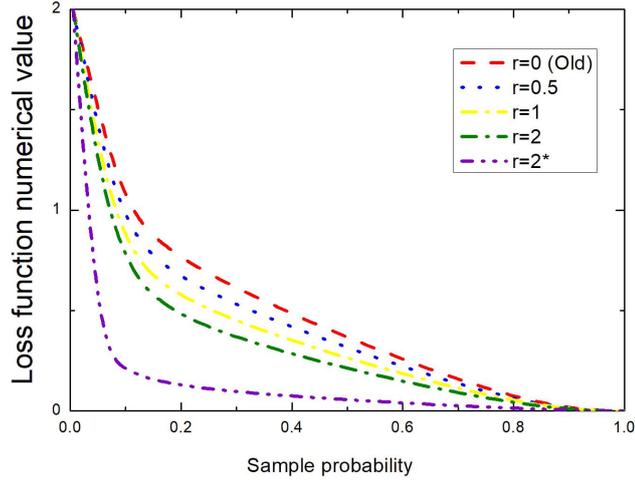
383

384 *Improvement based on Focal Loss*

385 The main reason that single-level detection is not as accurate as two-level detection is the
 386 imbalance of sample categories. Category imbalance will bring too many negative samples, which
 387 account for most of the loss function. Therefore, the focal loss is proposed as a new loss function.

388 The loss function is modified based on standard cross entropy loss in Figure 6. This function can
 389 reduce easily classified samples by changing the evaluation method, so that more weight is
 390 applied to hard classified samples in the training process [9]. The formula is as follows:

$$391 \quad FL(p_t) = - \alpha t(1 - p_t)^Y \log(p_t) \quad (21)$$



392

393 **Figure 6.** schematic diagram of loss function.

394 Firstly, a factor is added to the original standard cross entropy loss, thereby reducing the loss of
 395 easily classified samples. This makes us pay more attention to difficult and misclassified samples.
 396 For example, $\gamma=2$, for a positive sample with a prediction result of 0.95, the loss function value
 397 becomes smaller because the power of $(1-0.95)$ is small. However, for negative samples with a
 398 prediction probability of 0.3, the loss becomes relatively large, which is achieved by suppressing
 399 the loss of positive samples.

400 Therefore, the new method pays more attention to this indistinguishable sample. In this way, the
 401 influence of simple samples is reduced, and the effect will be more effective only when a large
 402 number of samples with small prediction probability are added together. Meanwhile more
 403 punishments are required for negative samples that are easily divided. The actual formula is as
 404 follows:

$$405 \quad L_{fl} = \begin{cases} -\alpha(1-y)^\gamma \log y' & y = \text{positive simple} \\ -(1-\alpha)y^\gamma \log(1-y'), & y = \text{negative simple} \end{cases} \quad (22)$$

406 In the experiment, $\gamma=2$ and $\alpha=0.25$ have the best effect.

407 *Improvement based on KL loss*

408 The traditional boundary box regression loss (i.e., Smooth $L1$ loss) does not take the deviation of
 409 the actual boundary on the ground into consideration. When the classification score is very high,
 410 the regression function is considered accurate, but it is not always the case.

411 Bounding box prediction is modeled as a Gaussian distribution, and the bounding box of positive
 412 samples is modeled as a Dirac delta function. The asymmetry of these two distributions is
 413 measured by KL divergence. When KL divergence approaches 0, these two distributions are very
 414 similar. KL loss is the KL divergence of minimizing the Gaussian distribution predicted by the
 415 bounding box and the Dirac delta distribution of positive samples. In other words, KL loss makes
 416 the bounding box prediction approximate to the Gaussian distribution, close to the positive sample.

417 And it converts the confidence into the standard deviation of the bounding box prediction.

418 The two probability distributions P and Q of a discrete or continuous random variable whose KL
 419 divergence is defined as:

$$420 \quad D(P\|Q) = \sum_{i \in X} P(i) * \left[\log \left(\frac{P(i)}{Q(i)} \right) \right] \quad (23)$$

$$421 \quad D(P\|Q) = \int_x P(x) * \left[\log \left(\frac{P(x)}{Q(x)} \right) \right] dx \quad (24)$$

422 Before calculating KL divergence, the bounding box needs to be parameterized. (x_1, y_1, x_2, y_2) is
 423 the upper left and lower right coordinates of the prediction bounding box. $(x_1^*, y_1^*, x_2^*, y_2^*)$ is the
 424 coordinates of the upper left and lower right corners of the real box. $(x_{1a}, y_{1a}, x_{2a}, y_{2a}, h_a, w_a)$ is an
 425 anchor bounding box generated by clustering all real boxes. Then the deviations of the predicted
 426 and real bounding boxes from are as follows:

$$427 \quad t_{x1} = \frac{x_1 - x_{1a}}{w_a}, t_{x2} = \frac{x_2 - x_{2a}}{w_a} \quad (25)$$

$$428 \quad t_{y1} = \frac{y_1 - y_{1a}}{h_a}, t_{y2} = \frac{y_2 - y_{2a}}{h_a} \quad (26)$$

$$429 \quad t_{x1}^* = \frac{x_1^* - x_{1a}}{w_a}, t_{x2}^* = \frac{x_2^* - x_{2a}}{w_a} \quad (27)$$

$$430 \quad t_{y1}^* = \frac{y_1^* - y_{1a}}{h_a}, t_{y2}^* = \frac{y_2^* - y_{2a}}{h_a} \quad (28)$$

431 Similarly, the parameter without * indicates the deviation between the prediction and the anchor
 432 boundary frame, and the parameter with * indicates the deviation between the real and the anchor
 433 boundary frame.

434 Assuming that the coordinates are independent, a univariate Gaussian function is used for
 435 simplicity:

$$436 \quad P_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_e)^2}{2\sigma^2}} \quad (29)$$

437 Where x_e is the estimated boundary box position and standard deviation σ is the estimated
 438 uncertainty. When $\sigma \rightarrow 0$, the boundary box position accuracy is very high.

439 The real boundary box on the ground can also be expressed by Gaussian distribution, and becomes
 440 Dirac delta function when $\sigma \rightarrow 0$:

$$441 \quad P_D(x) = \delta(x - x_g) \quad (30)$$

442 Where x_g is the real boundary box position on the ground.

443 At this time, we can construct a bounding box regression function with KL loss, and establish a
 444 formula to minimize the KL error of $p_\theta(x)$ and $P_D(x)$ on n samples:

445
$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \frac{1}{N} \sum D_{KL}(P_D(x) \| P_{\Theta}(x)) \quad (31)$$

446 KL divergence is used as the loss function L_{reg} for bounding box regression, and the classification
 447 loss L_{cls} remains unchanged. For a single sample:

$$\begin{aligned}
 L_{reg} &= D_{KL}(P_D(x) \| P_{\Theta}(x)) \\
 &= \int P_D(x) \log P_D(x) dx - \int P_D(x) \log P_{\Theta}(x) dx \\
 &= \frac{(x_g - x_e)^2}{2\sigma^2} + \frac{\log(\sigma^2)}{2} + \frac{\log(2\pi)}{2} - H(P_D(x))
 \end{aligned}$$

448 (32)

449 When the prediction of the bounding box is inaccurate, because the prediction closer to the real
 450 bounding box is certainly stable and its variance is small, the smallest possible variance can
 451 reduce Lreg. After the variance of the predicted positions of the bounding boxes is obtained, the
 452 candidate positions are voted according to the known variance of adjacent bounding boxes.
 453 Besides, the candidate coordinate values with the largest score are selected to be weighted to
 454 update the coordinates of the bounding box so as to make the positioning more accurate. What's
 455 more, border boxes with lower positions and lower colors have higher weights. The new
 456 coordinates are calculated as follows:

$$\begin{aligned}
 p_i &= e^{-(1-IoU(b_i,b))^2/\sigma_t} \\
 x &= \frac{\sum_i p_i x_i / \sigma_{x,i}^2}{\sum_i p_i / \sigma_{x,i}^2} \quad \text{subject to } IoU(b_i,b) > 0
 \end{aligned}$$

457 (33)

458 Where σ_t is an adjustable parameter for variable voting. When $IoU(b_i,b)$ is larger, p_i is larger,
 459 that is, the two bounding boxes overlap each other more and do the same for the remaining
 460 coordinate values. SSD detects the generated preselected box computing loss through FL loss
 461 function classification and border regression. Besides, the border regression of SSD is improved
 462 based on KL loss method. Frames with large variance and adjacent border frames containing the
 463 selected frames but too small will get low scores when voting. Moreover, the SSD algorithm can
 464 effectively avoid the abnormal situation mentioned above by variance voting instead of IoU
 465 overlap degree.

466 Model testing and analysis

467 The environment perception is divided into two parts, the micro part is main perception of the
 468 scene by machine vision, which is used to confirm and supplement the macrocosm and microcosm
 469 perception.

470 First of all, we tested the *Roi* weighting using live campus photos taken on May 7, 2020. The
 471 advantage of this algorithm is that the region of interest can be identified first, and then the further
 472 perception can be completed. Therefore, the test of region of interest was performed first, and the
 473 effect of attention weighting was significant.

474 Second, the environment perception test was carried out because the region of interest was

475 weighted and the weighted region was described firstly. After testing, the algorithm can complete
 476 the perception of the simple traffic scene and recognize the red light of the intersection, the bus
 477 and the right-turn sign on the road, and can supplement and confirm the environment perception
 478 part.

479 At the same time, different databases, Google Nic, Log BILINEAR and other different algorithms
 480 are used to compare with experiments, because the Algorithm has good performance on Flickr8K,
 481 Flickr30K and MS COCO databases , and validated the experimental results of the Northwestern
 482 Polytechnical University team. The experimental results on the Flickr8K database are shown in
 483 Table 5, on the Flickr30K database are shown in Table 6, and on the MS COCO database are
 484 shown in Table 7.

Ways	BLEU-1	BLEU-2	BLEU-3	BLEU-4
Google NIC	63	41	27	-
Log Bilinear	65.6	42.4	27.7	17.7
The design algorithm	67.2	45.8	32.1	18.9

485 Table 5. The results of the experiments on the database Flickr8K.

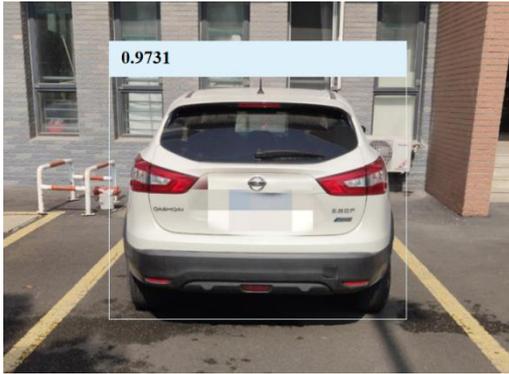
Ways	BLEU-1	BLEU-2	BLEU-3	BLEU-4
Google NIC	66.3	42.3	27.7	18.3
Log Bilinear	60	38	25.4	17.1
The design algorithm	67.1	45.1	29.9	21.1

486 Table 6. The results of the experiments on the database Flickr30K.

Ways	BLEU-1	BLEU-2	BLEU-3	BLEU-4
Google NIC	66.6	46.1	32.9	24.6
Log Bilinear	70.8	48.9	34.4	24.3
The design algorithm	72.3	51.8	37.1	25.1

487 Table 7. The results of the experiments on the database MS COCO.

488 The focus will be on target detection in the hazard prediction section. First of all, the vehicles test
 489 is carried out, using field test maps and dataset pictures. Secondly, dynamic vehicles need to be
 490 detected, including their speed, distance and running direction. The vehicle target detection is
 491 shown in Figure 7 and Figure 8. The dynamic vehicle direction estimation is shown in Figure 9
 492 and Figure 10. The dynamic vehicle distance estimation is shown in Figure 11. The vehicle speed
 493 detector is used to detect the speed of dynamic vehicles in Figure 12.



494



495 **Figure 7.** Vehicle target detection.

Figure 8. Vehicle target detection.

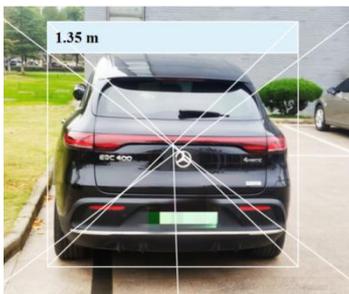


496



497 **Figure 9.** Dynamic vehicle direction estimation.

Figure 10. Dynamic vehicle direction estimation.



498



499 **Figure 11.** Dynamic vehicle distance estimation.

Figure 12. Vehicle speed detector.

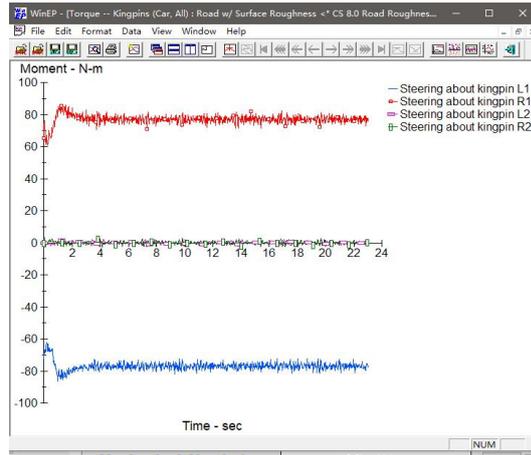
500 **Results and discussion**

501 **Simulated route**

502 In order to better reflect function of the system, the paper uses Matlab and Carsim to set up the
 503 dangerous situation of vehicle crossing in different conditions and conduct a joint simulation. The
 504 simulation system will output the speed constraint throughout the whole simulated driving process.
 505 Figure 13(a) simulates the vehicle operation by adjusting the scene, road surface and definition,
 506 driving conditions, etc. Figure 13(b) shows the speed constraint of the simulation system output in
 507 the whole simulation driving process.



(a) User Interface



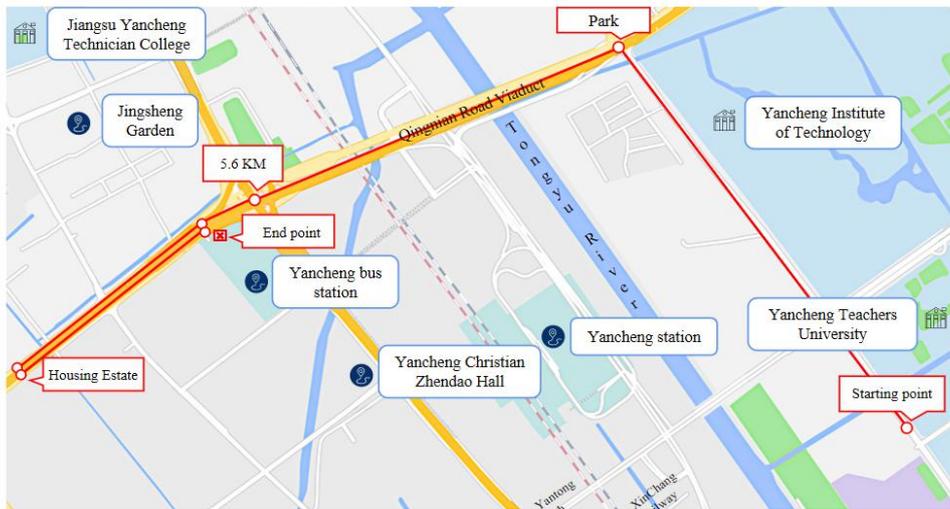
(b) Waveform Interface

508

509

510 **Figure 13.** Simulation interface of CarSim.

511 Choose the route from school to the bus station. The path passes through two campuses, two
 512 residential areas, a commercial center and four intersections. The total length of this path is 5.6
 513 kilometers, which can meet the needs of system test and simulation. In order to facilitate the
 514 simulation, the path of latitude and longitude are sampled. In addition, the path can be divided into
 515 two paths, Tianshan Road and Qingnian road, and the results are shown in Table 8 and Table 9
 516 respectively. The system test path map is shown in Figure 14.



517

518 **Figure 14.** Simulated route.

Positions	Latitude and longitude coordinates	Remarks
Starting point	120.205976,33.37899	West Gate of Yancheng Normal University's New Long Campus
Tianshan road	120.20524,33.37991	-

Crossroads	120.204431,33.380875	Century avenue and Tianshan Road Intersect
Tianshan road	120.202715,33.382887	-
Yancheng institute of technology	120.201997,33.38361	-
Tianshan road	120.200865,33.384786	-
Tianshan road	120.198403,33.387544	Unknown path junction
Xinfeng community	120.198403,33.387544	-
Crossroads	120.195753,33.390318	Tianshan Road and Youth Road Intersect

519 **Table 8.** Latitude and longitude sampling table of Tianshan road.

Positions	Latitude and longitude coordinates	Remarks
Youth road	120.193714,33.389948	-
Tongyu River Bridge	120.190229,33.388803	-
Tongyu River Bridge	120.186968,33.38768	There is a side road junction.
Youth road	120.181596,33.385909	There is a side road junction.
Complex intersection	120.180096,33.385344	Youth Road and Fangong Road Intersect
Howard square	120.174985,33.382133	-
Crossroads	120.173907,33.381576	Youth Road and Wengang Road Intersect
Youth road	120.176206,33.382812	-
End	120.178425,33.384033	Tianshan Road and Youth Road Intersect

520 **Table 9.** Longitude and latitude sampling table of Qingnian road.

521 Latitude and longitude sampling table of Tianshan road is shown in Table 10. Longitude and
522 latitude sampling table of Qingnian road is shown in Table 11. Based on the collected coordinates
523 of latitude and longitude, the whole system can be simulated and tested. In Matlab, the REGEXP
524 function can be used to get a Web page, so as to get the location name directly through the map
525 API and the output environment data through Json. Then the starting latitude and longitude for the
526 test are selected to successfully obtain the remote data.

527 System Simulation

528 In order to facilitate the simulation of the system function, the speed constraint on the simulation
 529 path is visualized. Considering the unity of safety and efficiency, the time has a great influence on
 530 speed constraint. Assuming that the vehicle is traveling at 60km/h, simulation speed constraints
 531 are provided for Monday, Sunday at 8:00 and Monday at 8:00 and 23:00. By the way, the system
 532 will adjust appropriately according to the risk weighting.

533 The speed constraint was loaded into Carsim for dynamic simulation, and the data at 23:00 on
 534 Monday was selected to check the difference between simulations with and without the system.

Positions	Latitude and longitude coordinates	Remarks
Starting point	120.205976,33.37899	West Gate of Yancheng Normal University's New Long Campus
Tianshan road	120.20524,33.37991	-
Crossroads	120.204431,33.380875	Century avenue and Tianshan Road Intersect
Tianshan road	120.202715,33.382887	-
Yancheng institute of technology	120.201997,33.38361	-
Tianshan road	120.200865,33.384786	-
Tianshan road	120.198403,33.387544	Unknown path junction
Xinfeng community	120.198403,33.387544	-
Crossroads	120.195753,33.390318	Tianshan Road and Youth Road Intersect

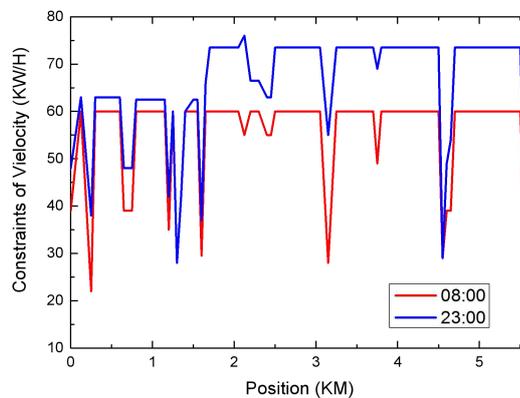
535 **Table 10.** Latitude and longitude sampling table of Tianshan road.

Positions	Latitude and longitude coordinates	Remarks
Youth road	120.193714,33.389948	-
Tongyu River Bridge	120.190229,33.388803	-
Tongyu River Bridge	120.186968,33.38768	There is a side road junction.
Youth road	120.181596,33.385909	There is a side road junction.
Complex intersection	120.180096,33.385344	Youth Road and Fangong Road Intersect
Howard square	120.174985,33.382133	-

Crossroads	120.173907,33.381576	Youth Road and Wengang Road Intersect
Youth road	120.176206,33.382812	-
End	120.178425,33.384033	Tianshan Road and Youth Road Intersect

536 **Table 11.** Longitude and latitude sampling table of Qingnian road.

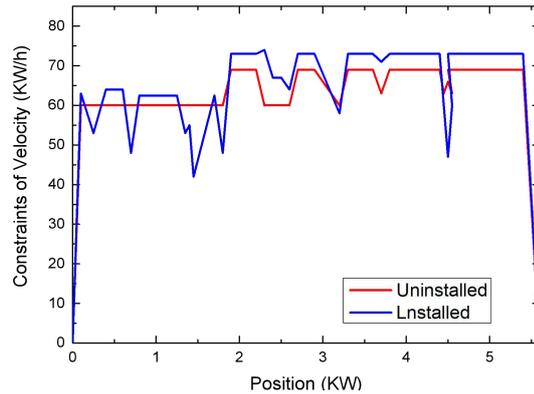
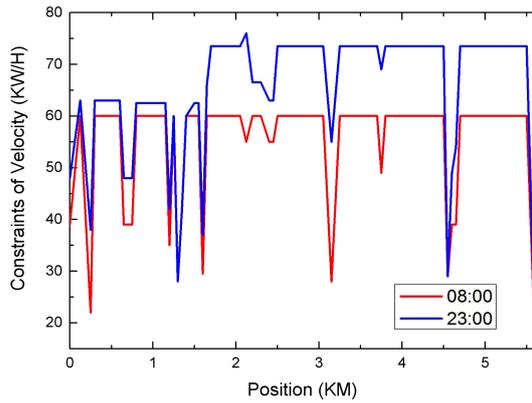
537 What's more, the system will adjust appropriately according to the risk weighting. The speed
538 constraint was loaded into Carsim for dynamic simulation, and the data at 23:00 on Monday was
539 selected to check the difference between simulations with and without the system.



540

541 **Figure 15.** Speed constraints at 8:00 on Monday and Sunday.

542 The comparison of speed constraints between Monday and Sunday is shown in Figure 15. On the
543 whole, the speed constraints on Monday are stricter than those on Sunday, which is caused by risk
544 weighting based on experience. And on the basis of time weighting, roads and intersections of
545 different levels are weighted simultaneously by the system, and corresponding speed constraints
546 are finally formed. At this time, the speed constraint does not consider the vehicle dynamics or the
547 comfort of driving. In practical application, the speed constraint needs to consider the required
548 acceleration of the current speed of the vehicle to implement the speed constraint. The acceleration
549 needs to be comprehensively considered according to the center of gravity, braking performance,
550 acceleration performance, ground friction coefficient, etc., which are ignored during the
551 simulation.



552

553 **Figure 16.** Speed Constraints at 8:00 and 23:00 on Monday. **Figure 17.** Velocity comparison.

554 The comparison results of speed constraints at different times on the same day are shown in Figure
 555 16. The speed constraint at 23:00 has been relaxed and vehicles are allowed to travel beyond the
 556 standard speed. In practical application, the unmanned driving system needs to combine the road
 557 supervision situation with the traffic situation on site to execute the speed. This speed only outputs
 558 speed constraints from the perspective of environmental hazards and does not represent the final
 559 execution speed.

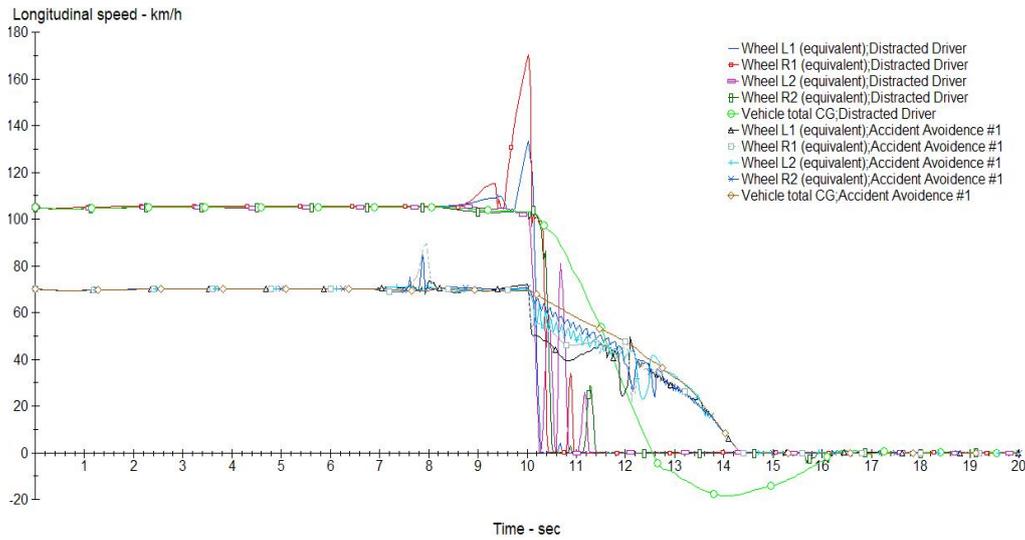
560 The comparison of simulated driving speeds of vehicles equipped with the system and
 561 human-driven vehicles is shown in Figure 17. Since most of the front part of the simulated route
 562 passes schools and intersections while other parts are expressways with few intersections,
 563 different driving speeds are simulated on the basis of actual driving. Under the ideal condition of
 564 smooth traffic, human driving vehicles will be affected by road grade, traffic control and
 565 subjective judgment. Besides, the driving speed of vehicles equipped with this system is similar to
 566 that of human beings in trend, and the speed constraint is strictly implemented according to the
 567 risk grade. It can be seen that the vehicles can realize the defensive driving of human beings more
 568 intelligently and flexibly, relying on accurate scientific and objective data analysis conclusions
 569 instead of subjective experience.

570 In order to reflect the efficiency of the system more particularly, the application test of the vehicle
 571 with or without the system is carried out through the car accident simulation built in CarSim.

572 The content of the traffic accident simulation is that an oncoming vehicle with a speed exceeding
 573 100km/h strays into the lane while avoiding the normally running vehicle and eventually rolls over.
 574 The normally running vehicle with a speed of 100km/h completes emergency braking during the
 575 avoidance. The accident site is a freeway.

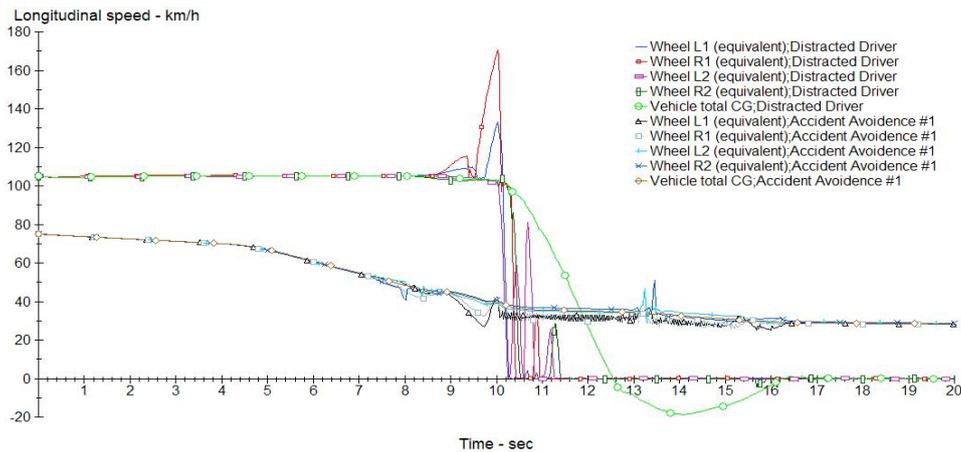
576 In the simulation, the speed of 70km/h is set as the normal driving speed, which is consistent with
 577 the actual use of expressway. Figure 18(a) demonstrates that the unsystematic vehicles are running
 578 normally at a speed of 70km/h. Within a period of visible sight distance, no vehicles were
 579 observed to enter the lane in opposite direction. After the danger was found, the collision was
 580 avoided by emergency braking. Figure 18(a) shows the speed of each tire of both vehicles. It
 581 indicates that the speed of the vehicles decreased sharply and fluctuated greatly during emergency

582 braking. The whole braking process has a great influence on passengers and has great uncertainty
583 factors.



584

585 (a) Simulation curve of vehicle avoidance without the system



586

587 (b) Simulation curve of vehicle avoidance with the system

588 **Figure 18.** CarSim Accident Simulation Diagram.

589 Figure 18(b) shows that the vehicles equipped with the system decelerated before the intersection
590 due to speed constraints, and drove through the intersection at a speed approaching 40km/h. After
591 the other vehicle entered the sight distance, lane change and braking operations were adopted to
592 avoid collision safely and stably. It can be seen that the braking curve of the vehicle with a system
593 is smoother than that of the vehicle without a system. The danger has been successfully avoided
594 with preparation in advance. Therefore, the simulation application proves the effectiveness of the
595 system to some extent.

596 Conclusion

597 Environmental hazard prediction and avoidance technology is the key in the research field of

598 unmanned vehicles, which provides an important guarantee for the driving of unmanned vehicles
599 in the real environment. This technology is of great significance to the improvement of the safety
600 and the commercialization of unmanned vehicles. Nowadays, most unmanned driving systems are
601 equipped with hazard prediction and avoidance systems. However, environment-oriented
602 data-based environmental hazard prediction and avoidance technology has not been developed.
603 Therefore, the design of this system has an advanced vision and provides a stronger guarantee for
604 the safety of unmanned driving. In this paper, Matlab and CarSim are used to simulate the entire
605 system. The speed constraint and simulation speed diagrams under various conditions are output on
606 the selected driving path to verify the effectiveness of the system function. And an accident is
607 simulated, which reflects the value of the system to a certain extent.

608 The system is innovative in the target detection of core technologies. In the overall design, guided
609 by deep learning and by the industry frontiers such as recognizing 5G and big data, it aims to
610 solve the problems of the unmanned industry based on the existing technological achievements
611 and realize the application transformation providing a reference to the improvement of the
612 unmanned industry.

613 **References**

- 614 1. X.Y. Z. et al. A study on key technologies of unmanned driving. *CAAI Transactions on*
615 *Intelligence Technology*.**1**, 4-13 (2016).
- 616 2. B. K. et al. Automated Complex Urban Driving based on Enhanced Environment
617 Representation with GPS/map, Radar, Lidar and Vision. *8th IFAC -AAC-II*. vol. **49** (IFAC,
618 2016).
- 619 3. Y.J. Li. et al. Research on static decoupling algorithm for piezoelectric six axis force/torque
620 sensor based on LSSVR fusion algorithm. *Mech Syst Signal Pr.* **110**, 509-520 (2018).
- 621 4. R. Z. et al. Electric vehicles' energy consumption estimation with real driving condition data
622 *Transport Res D-Tr E.* **41**, 177-187 (2015).
- 623 5. Y, C. et al. Landslide detection of hyperspectral remote sensing data based on deep learning
624 with constraints. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **12**, 5047–5060 (2019).
- 625 6. J.J. L. et al. A novel multi-segment feature fusion based fault classification approach for
626 rotating machinery. *Mech Syst Signal Pr.* **122**, 19-41 (2019).
- 627 7. F. W. et al. A novel integrated approach for path following and directional stability control
628 of road vehicles after a tire blow-out. *Mech Syst Signal Pr.* **93**, 431-444 (2017).
- 629 8. Khalid bin Hasnana. et al. JOMS: System Architecture for Telemetry and Visualization on
630 Unmanned Vehicle. *Procedia Engineering.* **29**, 3899-3903 (2012).

- 631 9. Z.Z. Li. et al. Influence of distance from traffic sounds on physiological indicators and
632 subjective evaluation. *Transport Res D-Tr E.* **87**, 102538 (2020).
- 633 10. Benjamin Ryder. et al. Spatial prediction of traffic accidents with critical driving events -
634 Insights from a nationwide field study. *Transport Res A-Pol.* **124**, 611-626(2019).
- 635 11. Hiroshi Yoshitake, MotokiShino. Risk assessment based on driving behavior for preventing
636 collisions with pedestrians when making across-traffic turns at intersections. *IATSS Research.*
637 **42**, 240-247(2018).
- 638 12. Rawia Ahmed Hassan E.L. Rashidya. et al. A big data modeling approach with graph
639 databases for SPAD risk. *Safety Sci.***110**, 75-79 (2018).
- 640 13. Brian S. Freeman. et al. Vehicle stacking estimation at signalized intersections with unmanned
641 aerial systems. *IJTST.* **8**, 231-249 (2019).
- 642 14. Y.W. C. et al. An effective infrared small target detection method based on the human
643 visual attention. *Infrared Phys Techn.* **95**, 128-135 (2018).
- 644 15. Debojit Biswas. et al. An automatic traffic density estimation using Single Shot Detection
645 (SSD) and MobileNet-SSD. *Physics and Chemistry of the Earth, Parts A/B/C.* **110**, 509-520
646 (2019).
- 647 16. A.G. et al. Deep learning for object detection and scene perception in self-driving cars:
648 Survey, challenges, and open issues. *Array.* **10**, 100057 (2021).
- 649 17. Farid Feyzi. et al. FPA-FL: Incorporating static fault-proneness analysis into statistical fault
650 localization. *J Syst Software.* **136**, 39-58 (2018).
- 651 18. Jan Březina. et al. Fast algorithms for intersection of non-matching grids using Plücker
652 coordinates. *Comput Mathe Appl.***74**, 174-187 (2017) .
- 653 19. Rondell Burge. et al. An investigation of the effect of texting on hazard perception using
654 fuzzy signal detection theory (fSDT). *Transport Res F-Traf.* **58**, 123-132 (2018).
- 655 20. Ellen F. Grumert. et al. Using connected vehicles in a variable speed limit system.
656 *Transport Res Procedia.***27**, 85-92 (2017).
- 657 21. El-Gamal, A. & Saleh, I. Radiological and mineralogical investigation of accretion and
658 erosion coastal sediments in Nile Delta region, Egypt. *J. Oceanogr. Mar. Sci.* **3**, 41-55 (2012).
- 659 22. Zhang, C. et al. Identifying and mapping individual plants in a highly diverse high-elevation
660 ecosystem using UAV imagery and deep learning. *ISPRS J. Photogramm. Remote. Sens.* **169**,
661 280–291 (2020).
- 662 23. Zhu, X. X. et al. Deep learning in remote sensing: a comprehensive review and list of
663 resources. *IEEE Geosci. Remote Sens. Mag.* **5**, 8–36 (2017).
- 664 24. Cui, B., Fei, D., Shao, G., Lu, Y. & Chu, J. Extracting Raft aquaculture areas from remote

665 sensing images via an improved U-Net with a PSE structure. *Remote Sens. Basel.* 11, 2053
666 (2019).
667 25. Y.T. B. et al. Crash prediction with behavioral and physiological features for advanced vehicle
668 collision avoidance system. *Transport Res C-Emer.* 74, 22-23(2017).

669 **Acknowledgements**

670 This work is supported by the National High Tech Project "863" [Project No. 2011AA11A286] and
671 Natural Science Foundation of China [Project No. 51405201].

672 **Author contributions**

673 C.Q.Q conducted experimental work; drafted manuscript; analysed data. Y.Z. assisted with
674 measurements. J.J., S.Z., H.Z.,S.Q.Z., assisted with target detection and analysis of data.
675 M.Y.M.supervised and conceived study. All authors commented on and reviewed paper.

676 **Competing interests**

677 The authors declare no competing interests.