



Article

# Quantifying the Effects of Visual Road Information on Drivers' Speed Choices to Promote Self-Explaining Roads

Yuting Qin <sup>1</sup> , Yuren Chen <sup>1,2,\*</sup> and Kunhui Lin <sup>1</sup>

<sup>1</sup> The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Shanghai 201804, China; 1833257@tongji.edu.cn (Y.Q.); 1731276@tongji.edu.cn (K.L.)

<sup>2</sup> Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 201804, China

\* Correspondence: chenyr@tongji.edu.cn; Tel.: +86-137-0166-4204

Received: 9 March 2020; Accepted: 1 April 2020; Published: 3 April 2020



**Abstract:** Roads should deliver appropriate information to drivers and thus induce safer driving behavior. This concept is also known as “self-explaining roads” (SERs). Previous studies have demonstrated that understanding how road characteristics affect drivers' speed choices is the key to SERs. Thus, in order to reduce traffic casualties via engineering methods, this study aimed to establish a speed decision model based on visual road information and to propose an innovative method of SER design. It was assumed that driving speed is determined by road geometry and modified by the environment. Lane fitting and image semantic segmentation techniques were used to extract road features. Field experiments were conducted in Tibet, China, and 1375 typical road scenarios were picked out. By controlling variables, the driving speed stimulated by each piece of information was evaluated. Prediction models for geometry-determined speed and environment-modified speed were built using the random forest algorithm and convolutional neural network. Results showed that the curvature of the right boundary in “near scene” and “middle scene”, and the density of roadside greenery and residences play an important role in regulating driving speed. The findings of this research could provide qualitative and quantitative suggestions for the optimization of road design that would guide drivers to choose more reasonable driving speeds.

**Keywords:** road characteristics; speed choice; self-explaining road; random forest; convolutional neural network

## 1. Introduction

Speed is a crucial factor affected the occurrence and consequences of road traffic crashes [1,2]. Therefore, great importance should be assigned to speed management. There is no doubt that speed reduction measures such as speed limits and red pavement can help to ensure traffic safety. However, in many cases, when drivers fail to perceive risks and develop an inappropriate speed choice, they will “turn a blind eye” to these safeguards [3,4]. Thus, optimizing the layout of roads to improve drivers' risk perception and guide them towards adjusting their driving behaviors spontaneously is of great significance. When drivers' expectations coincide with the actual situation, driving will be much safer. The concept of road design based on human factors was put forward by Dutch scholar Theeuwes as early as 1995 [5], and is called self-explaining roads (SERs).

According to the international literature, clarifying the function and classification of roads is an effective measure that can be used to create SERs. This kind of approach focuses on psychological modeling of drivers' subjective perceptions of roads. The assumption is that drivers usually make judgments unconsciously, based on personality characteristics, existing knowledge, and driving experience [6–9]. Identifying the road features that drivers use to make discriminations is the key

to designing a uniform and predictable road system. Some explanations have been suggested as mechanisms of drivers' speed choice: (1) Abstract feelings like monotony, physical comfort, task difficulty, and safety are macro factors that determine drivers' subjective categorization of roads, among which comfort seems to be the most important [10,11]. (2) By visually inspecting clustered road scenarios, road width, lane markings, speed limits, and landscape as discriminant clues. A specific combination of these clues can successfully categorize roads in a way that relates to drivers' perceptions [12]. It is worth noting that road alignment, which has a great correlation with driving speed on rural roads, is much less important than land use for urban roads. Self-explaining roads have been constructed in New Zealand using these cognitive mechanisms. Observations have shown that by clarifying the classification of roads and expanding the differences between categories, the actual road accident rate is significantly reduced [13,14].

Quantifying the relationships between road characteristics and driving speed is another approach used in the design of SERs. In this approach, vehicle kinematic parameters such as driving speed and lateral deviation are regarded as a reflection of drivers' thoughts. Research methods used to measure their impact are summarized as follows: (1) Exploring differences in driving behavior under changeable road scenarios through simulated driving experiments [15,16]. By adjusting the target of the experiment, the influence and effectiveness of a specific road feature can be evaluated. (2) Establishing a behavior prediction model and assessing the importance of the predictors [17,18]. According to research findings worldwide, road attributes influencing driving speed can be enumerated: road alignment [19], roadside conditions [16], lane and shoulder width [20], speed limit [21], recovery-zone width, and junction density [22]. In China, operating speed models have been written into the specifications for highway safety audits to evaluate the consistency of road design. However, the models contain only geometric parameters such as horizontal radius and longitudinal slope.

In general, understanding the impact of road features on drivers' speed choices is the common theme of the above-mentioned research, and it is the key to SERs. Psychological research starts with the general layout of roads and focuses on the information that can be used to discriminate drivers' subjective judgments. However, research results are usually inferential and vague, and it is hard to generalize universally applicable engineering principles from that. In comparison, the quantitative method seems to be a better choice, but models developed to date vary greatly. Aside from the concept of methodology, the neglect of drivers' perceptions is another important factor. Although the studies mentioned above have some practical significance for promoting SERs, most of them include "improvement suggestions" but not "design". Extracting visual road information from images and calibrating its impact on driving speed were the objectives of this paper. Road images from the driver's visual perspective were separated into layers based on the image semantic segmentation technique. The influences of each layer on driving speed were quantified using speed tags rather than regression coefficients. This is an innovative and practical method of SER design. We hope that the research results can provide suggestions for road layout optimization using speed tags. When a road design needs to be adjusted, the appropriate road facilities can be indexed according to the expected speed tag.

The remaining parts of this paper are organized as follows. The next section introduces the source of the data and the methods of visual road information extraction. The main methodologies of this research are then introduced, including the method used to calculate the speed tag corresponding to each piece of information, the random forest algorithm, and the convolutional neural network. In the following part, speed prediction models were established using these two machine learning algorithms, and the importance of predictors was calculated. The last two sections provide a discussion of the research results and conclude with the contributions.

## 2. Experiments and Data

### 2.1. Naturalistic Driving Experiment

Naturalistic driving experiments were conducted on five two-lane rural roads in Tibet Province of China. Ten drivers (eight males and two females) took part in the experiment. Their ages ranged from 23 to 50 (mean = 32.9, std = 7.1). All of them had more than three years' driving experience. A driving recorder (GARMIN GDR35) was fixed on the windshield to obtain videos from the driver's visual point of view. A three-axis acceleration sensor was synchronized with the camera to gather kinematic vehicle information including driving speed, acceleration, and impulse forces. The total driving mileage was more than 800 km, and over 20 h of driving dash-cam video were obtained from the experiments. Road sections used in our study contained various traffic facilities and roadside landscapes. The traffic volume was relatively low. Given that the response time of most drivers is about 2 s [17], visual road scenarios were matched with driving speed 2 s after.

### 2.2. Visual Road Information Extraction

Maslow, an American psychologist, established the "hierarchy of needs" theory and studied human behaviors from the perspective of needs [23]. Considering that driving is also a demand-driven behavior, it can be assumed that driving speed is determined by road geometry and modified by the road environment. Road information perceived by drivers can also be described by reference to the hierarchy of requirements. Geometry information is essential and mandatory. In the process of driving, road alignment shows the extension of the road, and the driver has to follow it by controlling the steering wheel and pedal. Although lane markings sometimes do not exist, the driver can subliminally perceive the "shape" of the road and further perform the driving operations [18]. Environment information is additional and optional. For example, for a warning sign to successfully affect driving behavior, it must be detected, understood, and accepted [3]. Therefore, based on the necessity, the visual road information is classified and extracted as follows:

#### (1) Visual Geometry Information

According to our previous study, the Catmull–Rom spline can fit the road geometry from a driver's visual perspective well [24–26]. Since there is not always a centerline on rural roads, two Catmull–Rom splines were applied to fit the left and right lane boundaries. As shown in Figure 1c, the shape of each spline was controlled by four points ( $P_{Li}, P_{Ri}, I = 1, 2, 3, 4$ ) dividing the visual lane into "near scene", "middle scene", and "far scene". Boundary length and average curvature of each region were regarded as the shape parameters, and could describe the visual road geometry. They were calculated as follows.

$$vS_{Li(i+1)} = S_{Li+1} - S_{Li}, \quad (1)$$

$$vK_{Li(i+1)} = \frac{f_{Li+1} - f_{Li}}{vS_{Li(i+1)}}, \quad (2)$$

$$vS_{Ri(i+1)} = S_{Ri+1} - S_{Ri}, \quad (3)$$

$$vK_{Ri(i+1)} = \frac{f_{Ri+1} - f_{Ri}}{vS_{Ri(i+1)}}, \quad (4)$$

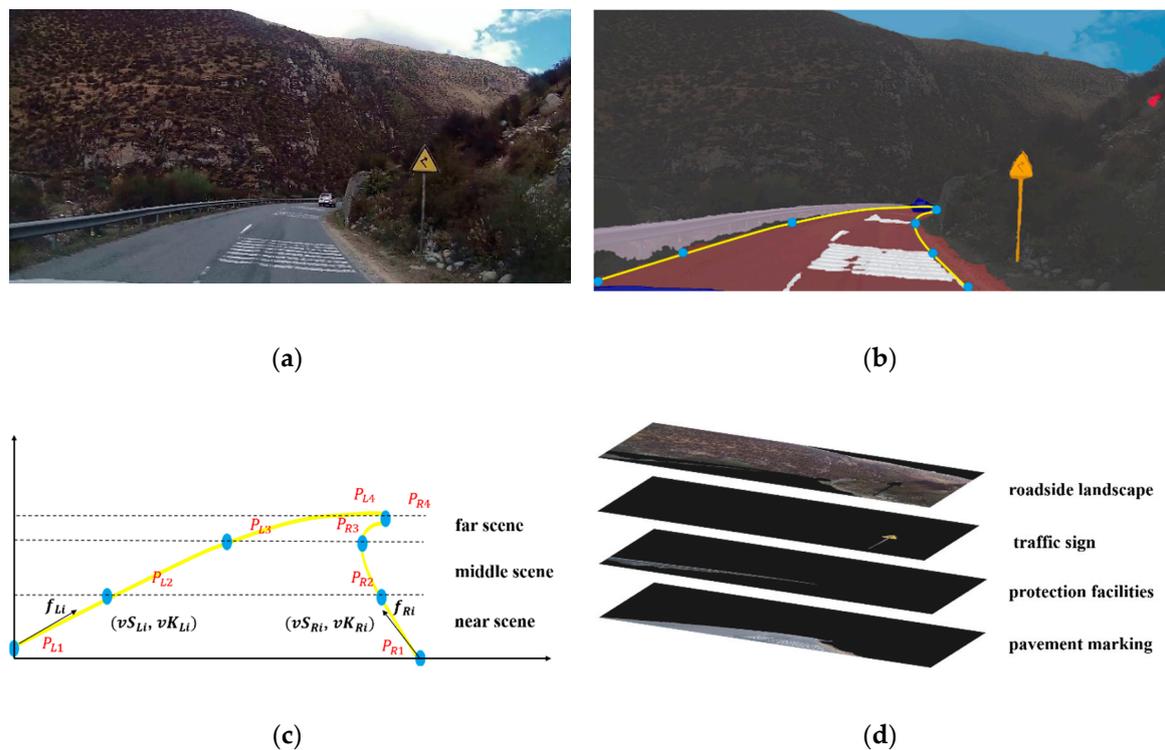
where  $i = 1, 2, 3$ ;  $vS_{Li(i+1)}$  is the boundary length between the control point  $P_{Li}$  and  $P_{Li+1}$  (measured in pixels).  $f_{Li+1}$  represents the tangent angle of  $P_{Li}$ ;  $vK_{Li(i+1)}$  is the average curvature between  $P_{Li}$  and  $P_{Li+1}$ .  $vS_{Ri(i+1)}$ ,  $f_{Ri+1}$ ,  $vK_{Ri(i+1)}$  are parameters for the right lane boundary, which are similar to those for the left one.

#### (2) Visual Environment Information

Compared with dummy variables, an RGB image, a high-dimensional matrix, can represent richer road environment information. Therefore, road images from the driver's perspective were

used to describe road environment information in this research. Since the disturbance of other vehicles, non-motor vehicles, pedestrians, and livestock was outside the scope of this research, objects observed by drivers were divided into four categories: roadside landscape (e.g., trees, mountains, buildings), traffic signs (e.g., speed limit, sharp curve warning), pavement markings (e.g., red pavement, transverse speed reduction markings), and protection facilities (e.g., guardrail, concrete barrier). An image semantic segmentation technique was used to layer the image. Based on a public dataset named Mapillary Vista, we reproduced the ICnet processed by Zhao et al. (for more information, please see Reference [27]). It performed well on our dataset. Pixels in the original image were classified into four categories and stored in different layers according to their category. Generally, up to four image layers could be separated from one road image. The size of all images was  $1920 \times 1080$ .

The process of visual information extraction is illustrated in Figure 1.



**Figure 1.** An instance of visual information extraction. (a) Original road scenario; (b) Lane boundary fitting and semantic segmentation; (c) Calculation of shape parameters; (d) Extraction of pixels by category.

### 3. Methodology

This methodology section consists of three parts: (1) Speed tags were identified by establishing a speed decision model. To calibrate the speed tags of each information, road scenarios were divided into three categories in terms of complexity. (2) The random forest algorithm and a convolutional neural network were used to predict the geometry-determined speed and environment-modified speed. In addition, methods used to measure variable importance in the corresponding models are described.

#### 3.1. Identification of Speed Tags

This study hypothesized that speed choice is ultimately determined by geometry information and modified by environment information. A speed decision model was put forward as follows.

$$V = V_g + \delta, \tag{5}$$

$$\delta = \sum_i^n \omega_i \delta_i, \quad (6)$$

where  $V$  is the final speed choice;  $V_g$  represents for the geometry-determined speed;  $\delta$  is the general correction caused by environment information; and  $n$  is the number of categories of environment information.  $i = 1, 2, 3, 4$ , which represent landscape, traffic sign, pavement markings and protection facilities, respectively.  $\delta_i$  stands for the speed change due to environment information, and  $\omega_i$  represents its weight. If  $\delta_i$  is negative, it indicates that the environment information has an inhibiting effect on the driver's expectation. When  $\delta_i$  is positive, it means that environment information promotes driver's expectations.

Road scenarios usually contain a variety of visual information, but there is only one driving speed. Therefore, controlling the variables is the only way to identify the speed tag of each piece of information. Geometry information exists in any road scenario, but the category of environment may include zero, one, or more pieces of information. Depending on the value of  $n$ , road scenarios were classified as crude scenarios, single-stimulus scenarios, or multi-stimulus scenarios, in which  $V_g$ ,  $\delta_i$ , and  $\omega_i$  were calibrated. The detailed method used was as follows.

- **In crude scenarios**, roads are laid out in an open field. There are no houses, greenery, or mountains on the side of the road. Traffic facilities do not exist either. Nothing but road geometry affects driving speed. In such road scenarios,  $V = V_g$ ,  $\delta = 0$ .
- **In single-stimulus scenarios**, there is one and only one category of environment information. In this case,  $n = 1$ ,  $\delta = \delta_i$ , and  $\omega_i = 1$ . By calculating  $V_g$  and  $\delta$ , the speed change stimulated by the existing environment information can be estimated.
- **In multi-stimulus scenarios**, there are more than two kinds of environment information. They compete for the driver's attention and affect driving speed collectively. On the basis of calculations in single-stimulus scenarios,  $\delta_i$  can be estimated. Statistical methods such as multiple linear regression and the Pearson correlation test were used to investigate  $\omega_i$ .

### 3.2. Machine Learning Algorithm

#### 3.2.1. Random Forest Algorithm

The random forest algorithm was used to build a regression model between the 12 visual shape parameters and  $V_g$ . It is an ensemble algorithm that integrates plenty of single regression trees (CART) by capturing their average of regression as the output. Bagging and boosting techniques are combined in this algorithm. Bagging is a method that can calculate many models at the same time, which can realize parallel computing and improve model robustness. Boosting is an approach for reducing bias. Outliers are highly tolerated, and the importance of explanatory variables can be evaluated in random forest models [28].

The Random Forest Regressor function in scikit-learn (a machine learning toolbox of python) was used to build the model. There are two hyperparameters to be regulated in this model, including the number of the trees in the forest ( $n_{trees}$ ) and the number of variables contained in each split ( $n_{features}$ ). About two-thirds of the samples were used to train the model, and the remaining one-third was used to evaluate the quality of the model [29]. Since driving speed is a continuous variable, MAE (mean absolute error), MSE (mean square error), and  $R^2$  (explained variance score) were adopted to assess the goodness-of-fit of the model. The grid research and cross-validation methods were utilized to find the optimal combination of the two hyperparameters [30]. The training process can be elaborated as follows.

- The number of trees in the forest was set to  $n_{trees}$ .
- A subset of the predictors was randomly selected as candidates for splitting, and the sampling size was equal to  $n_{features}$ .

- The best variable and split-point were picked out among the selected predictors, and each node was split into two subnodes.
- The output of every single tree was aggregated as the final output of the model.

With a pre-trained model, by adding random noise to a certain variable, the reduction of the predicting accuracy can be utilized to measure the relative importance of the predictors [31].

### 3.2.2. Convolutional Neural Network

A convolutional neural network (CNN) was trained to explore the relationship between  $\delta_i$  and layered images. CNNs are feedforward neural networks. They have exhibited excellent performance on image understanding due to their intelligent way of extracting critical features. Convolution and pooling are two important computing modules in most CNN models. The convolutional layers are designed to extract image features. The main function of pooling is down-sampling, which can remove redundant information. The bottom of a CNN is usually the fully connected (FC) layer. It can map the acquired image features to the outputs [32].

There are numerous popular CNN topologies, such as AlexNet [33], GoogleNet, and ResNet [34]. These models perform well in detail extraction, and additional training of these models for new images can save time and produce satisfying results [35]. As the pixels were filtered through semantic segmentation technique, a CNN with a relatively simple topology was developed for use in this study. According to our experience and trial calculation, a convolutional neural network of 10 layers was constructed. It consisted of four convolutional layers, four max pooling layers, and two fully connected layers. A rectified linear unit (ReLU) was inserted after each convolutional layer for non-linear activation. The architecture of the network is shown in Table 1.

**Table 1.** The architecture of the convolutional neural network.

Layer	Output Size	Number of Parameters
conv3-32, ReLU	(14814832)	896
2×2 max pool	(747432)	0
conv3-64, ReLU	(727264)	18496
2×2 max pool	(363664)	0
conv3-128, ReLU	(3434128)	73856
2×2 max pool	(1717128)	0
conv3-128, ReLU	(1515128)	147584
2×2 max pool	(77128)	0
FC-512, ReLU	(1512)	3211776
FC-3	(11)	513

To reduce the amount of computation required, images were resized from  $1920 \times 1080 \times 3$  to  $150 \times 150 \times 3$  before being entered into the network. Additionally, data augmentation techniques were utilized to improve the prediction accuracy and avoid overfitting; segmented images were randomly rotated (range from  $-10^\circ$  to  $10^\circ$ ), translated (within 20%), scaled (within 20%), and then inputted into the network. The goal of training was to minimize the MSE. A total 20% of the samples were randomly selected as the verification set, and the remaining 80% were considered the training set. The maximum training epoch was 500.

Neural networks perform well to solve non-linear problems, but the model is poorly interpretable. To observe the operation mechanism of the network, a class activation map (CAM) was proposed to visualize the calculation basis of the network. A CAM is a kind of heatmap that can visualize the points used by the model to make a particular decision by highlighting the determinative pixels. The importance value of each pixel can be calculated by multiplying the global average value of the gradients and feature maps obtained from the last convolution layer [36]. After the CNN model was trained, CAMs were drawn to visualize the important environmental information affecting the driving speed.

## 4. Results

### 4.1. Geometry-Determined Speed and Prediction Model

A total of 566 crude scenarios were picked out from the naturalistic driving experiment. The distribution of shape parameters and  $V_g$  values is shown in Table 2. From this, 378 samples were randomly selected to train the random forest regression model. The input of the model was the 12 visual shape parameters, and the output was  $V_g$ . In the process of grid researching, the MAE was found to decrease initially with increasing  $n_{trees}$ , and to level off when the value of  $n_{trees}$  exceeded 192. In most cases, with the same number of  $n_{features}$ , MAE was the minimum when  $n_{features} = 8$ . Therefore,  $n_{trees}$  was set to 192 and  $n_{features}$  was equal to 8. The prediction results are illustrated in Table 3. The final MAE was 1.29 and  $R^2$  was 0.96.

**Table 2.** Distribution of shape parameters and geometry-determined speed.

Parameter	Minimum	Maximum	Mean	Standard Deviation
$vK_{L12}$	$-8.91 \times 10^{-5}$	$1.29 \times 10^{-3}$	$8.89 \times 10^{-5}$	$1.74 \times 10^{-4}$
$vK_{L23}$	$-2.02 \times 10^{-2}$	$8.40 \times 10^{-3}$	$-2.95 \times 10^{-4}$	$4.02 \times 10^{-3}$
$vK_{L34}$	$-3.60 \times 10^{-2}$	$1.05 \times 10^{-2}$	$-4.19 \times 10^{-3}$	$8.17 \times 10^{-3}$
$vS_{L12}$	177.51	652.13	453.71	102.16
$vS_{L23}$	78.62	566.34	312.12	100.40
$vS_{L34}$	56.26	592.78	218.71	123.33
$vK_{R12}$	$-1.85 \times 10^{-3}$	$4.50 \times 10^{-4}$	$-1.04 \times 10^{-4}$	$2.63 \times 10^{-4}$
$vK_{R23}$	$-7.08 \times 10^{-3}$	$2.74 \times 10^{-2}$	$3.14 \times 10^{-3}$	$6.69 \times 10^{-3}$
$vK_{R34}$	$-1.57 \times 10^{-2}$	$3.40 \times 10^{-2}$	$1.05 \times 10^{-3}$	$7.42 \times 10^{-3}$
$vS_{R12}$	136.19	697.37	338.27	105.56
$vS_{R23}$	82.86	460.32	232.54	83.85
$vS_{R34}$	38.27	529.91	204.54	101.29
$V_g$	65.00	111.00	81.64	9.41

**Table 3.** Prediction result of random forest.

Training Set			Testing Set		
Sample Size	MAE	$R^2$	Sample Size	MAE	$R^2$
378	2.01	0.91	188	1.29	0.96

Variable importance was calculated and is illustrated in Figure 2. Curvatures of the right boundary in the “middle scene” and the “near scene” were the top-ranked two, with accuracy decreases of 33% and 26.82%. The following is the visual curve length of the left boundary. Interestingly, the curve length of the left boundary was more important than that of the right side. The curvature of the left boundary had less effect on driving speed. When noises were introduced into them, there was less than 4% reduction in prediction accuracy. The average curvature of the right boundary in the “far scene” was an unimportant predictor.

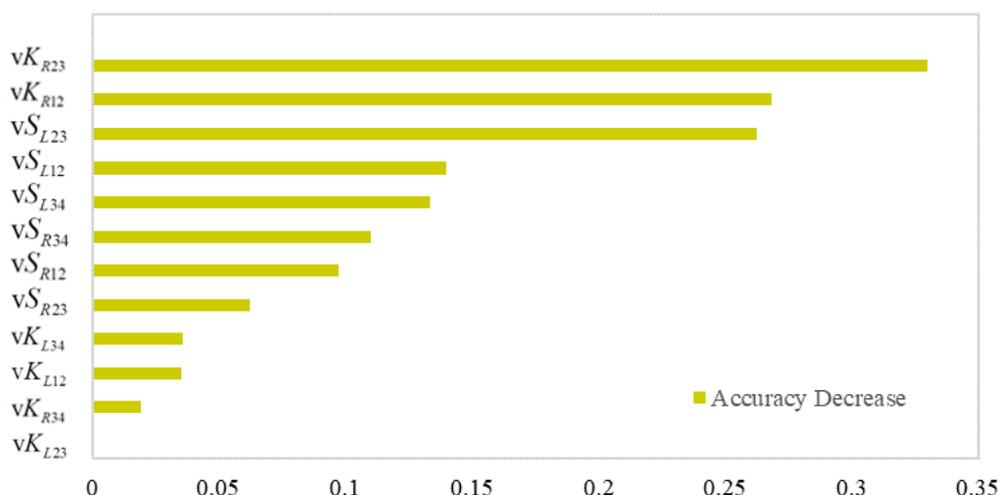


Figure 2. Variable importance in random forest.

4.2. Environment-Modified Speed and Prediction Model

A total of 623 single-stimulus road scenarios were found in the video data. According to the pre-trained geometry-determined speed model, the  $V_g$  of each scenario was computed. Statistical descriptions of  $V$ ,  $V_g$ , and  $\delta_i$  are shown in Table 4.  $V$  was generally smaller than  $V_g$  in single-stimulus scenarios, which indicates that environment information usually inhibited driving speed. Speed change caused by the landscape was relatively large and discrete. It was distributed within the range of (−50 km/h, 10 km/h), and two peaks appeared at −35 km/h and −5 km/h. It was inferred that there were two distinct kinds of landscapes, one with a large negative impact and one with a small negative impact on driving speed. The distributions of  $\delta_i$  caused by traffic signs and protection facilities were close to normal distributions, and the mean values of them were around −11 km/h and −14 km/h. The  $\delta_i$  caused by pavement markings was the smallest, while the standard deviation (std.) was the largest.

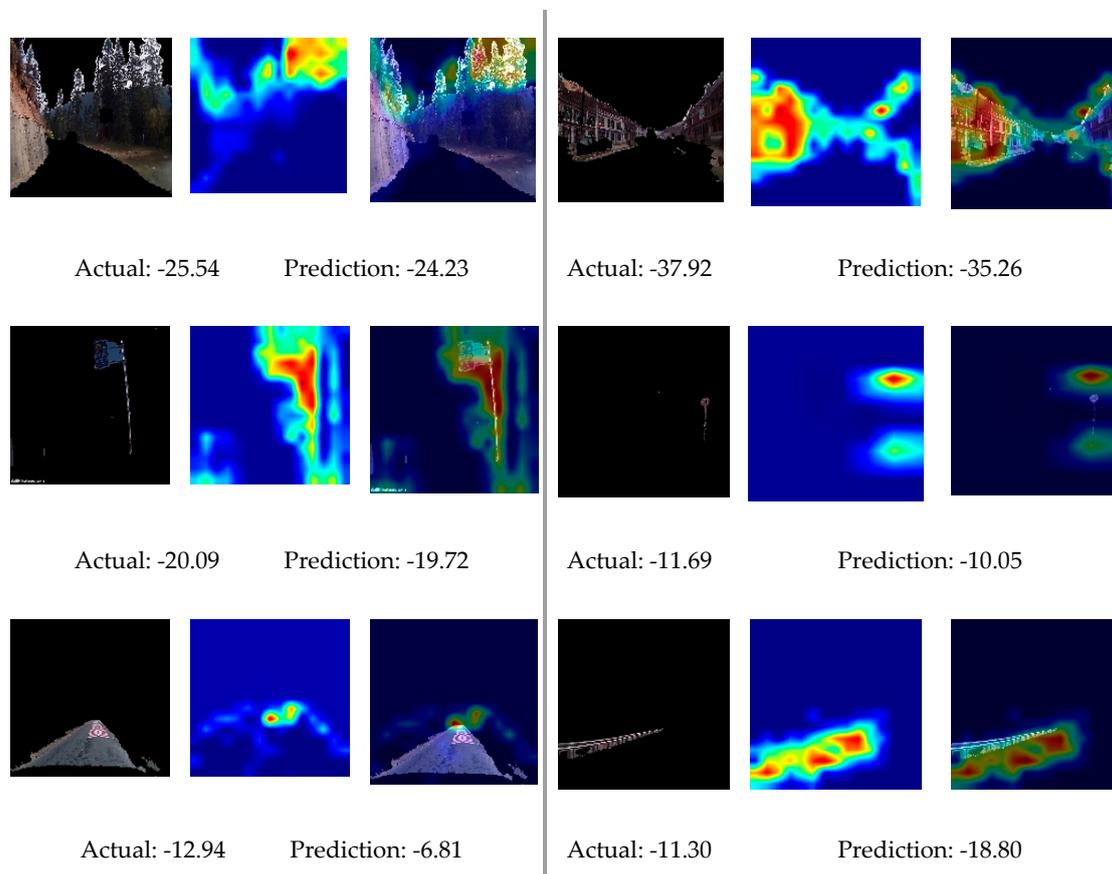
Table 4. The statistical descriptions of actual speed, geometry-determined speed and environment-modified speed.

Semantic Information Category	$V$		$V_g$		$\delta(\delta_i)$		Distribution of $\delta(\delta_i)$
	Mean	Std.	Mean	Std.	Mean	Std.	
Landscape	66.3	12.41	87.79	6.84	−26.23	11.26	
Traffic sign	75.69	11.83	87.86	4.96	−11.35	12.1	
Pavement marking	77.93	10.4	88.18	4.73	−9.7	12.63	
Protection facility	70.17	9.18	84.77	8.45	−13.8	9.13	

A convolutional neural network of 10 layers was constructed and trained with the layered image as the input and  $\delta_i$  as the output. The prediction result is shown in Table 5. It seemed that the CNN model learned to understand the environment images. To analysis the effectiveness of this model, CAMs were plotted and observed. Some examples, including the layered image, the CAM, and their superimposition, are presented in Figure 3. The key information needed to identify the value of  $\delta_i$  was concluded to be “treetop”, “windows”, and “position of the facility”.

**Table 5.** Training result of the convolutional neural network (CNN).

Training Set			Testing Set		
Sample Size	MSE	MAE	Sample Size	MSE	MAE
450	55.09	6.89	173	37.37	4.58



**Figure 3.** Recognition results from the CNN model.

### 4.3. Analysis of the Interaction of Environment Information

A total of 416 multi-stimulus scenarios were extracted from the video data.  $V_g$  and  $\delta$  were calculated first. By inputting layered images into the CNN model, speed changes caused by the landscape, traffic sign, lane markings and protection facilities (if present) were calculated and denoted as  $\delta_i$  ( $i = 1,2,3,4$ ). Multiple linear regression was performed to evaluate the relationship between  $\delta$  and  $\delta_i$ . The significance level was chosen to be 0.001, and the result is illustrated in Table 6. The model passed the F test ( $F(4403) = 9.471$ )  $p < 0.001$ , which indicated that the linear equation was statistically significant. However, only the  $\delta_i$  caused by landscape and pavement markings had a significant correlation with  $\delta$ . Additionally, the values of  $R$  and  $R^2$  were quite small, indicating that the goodness of fit and the variance interpretation ability of the model were not satisfactory. Therefore, the interaction of environment information was not a simple linear superposition, nor was  $\omega_i$  a fixed value.

The deviation between  $\delta$  and  $\delta_i$  can be used to roughly measure the relative importance of each piece of information because  $\delta$  is equal to  $\delta_i$  when there is no disturbance from other information. The greater the difference, the lower the relative importance. To investigate whether the type of environment information (represented by velocity correction capability  $\delta_i$ ) influenced the relative importance, the Pearson correlation coefficient between  $\delta_i$  and  $|\delta - \delta_i|$  was calculated. The results are shown in Table 7. The speed change caused by landscape had significant associations with the relative

importance of all categories of environment information. The correlation was mostly positive. The  $\delta_i$  caused by traffic signs had a significant and negative impact on the importance of landscape and itself, while the effect was positive for pavement markings. The  $\delta_i$  of pavement markings was positively associated with the effectiveness of the landscape, while the  $\delta_i$  of protection facilities was significantly and negatively related to the relative importance of itself.

**Table 6.** Results of multiple linear regression.

Model	Sum of Squares	ANOVA			F	Sig.
		df	Mean Square			
Regression	4720.78	4	1180.19	9.47	<0.001	
Residual	50,217.38	403	124.61			
Total	54,938.16	407				
Model Summary						
R	R <sup>2</sup>	Adjusted R <sup>2</sup>		Std. Error of the Estimate		
0.293	0.086	0.076856		11.16283		
Coefficients						
Variables	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
	B	Std. Error	Beta			
(Constant)	−12.852	2		−6.427	<0.001	
Landscape	0.224	0.054	0.209	4.128	<0.001	
Traffic sign	0.092	0.089	0.058	1.039	0.299	
Pavement marking	0.554	0.147	0.197	3.763	<0.001	
Protection facilities	−0.067	0.076	−0.049	−0.887	0.375	

**Table 7.** Result of Pearson correlation analysis.

Category			Relative Importance			
			Roadside Landscape	Traffic Sign	Pavement Markings	Protection Facilities
			Roadside landscape	Pearson correlation	−0.450 ** <sup>1</sup>	0.305 **
	Sig. (2-tailed)	0.000	0.001	0.008	0.000	
Traffic Signs	Pearson correlation	−0.323 **	−0.313 **	0.514 **	0.001	
	Sig. (2-tailed)	0.000	0.000	0.009	0.994	
Pavement Markings	Pearson correlation	0.332 **	0.379	0.089	0.156	
	Sig. (2-tailed)	0.000	0.062	0.316	0.227	
Protection Facilities	Pearson correlation	−0.001	−0.017	0.022	−0.555 **	
	Sig. (2-tailed)	0.993	0.895	0.864	0.000	

<sup>1</sup> \*\* represent that the correlation is significant at the 0.01 level (2-tailed).

## 5. Discussion

Previous studies have demonstrated the importance of drivers' visual perceptions [37], and some scholars have tried to extract color and shape information from road images [38] in order to interpret driving behaviors. In this research, road images were successfully mapped to driving speed based on machine learning methods. These results indicate that extracting visual road information from RGB images can effectively predict driving speed and reveal behavior mechanisms.

According to the random forest model, the curvatures of the right boundary in "near scene" and "middle scene" were the most important factors for drivers' speed choice. Similar conclusions were reported by Chen et al. [39] after analyzing data on drivers' eye movements on curved sections. The length of the left lane boundary, which reflects the sight distance to some extent, was also an important factor for driving speed. It is worth noting that the length of the boundary of the left lane was much more important than that of the right side. On curved sections, the visible distance of the left and right

lanes is usually different due to perspective transformation. Since the conflict with vehicles traveling in the opposite direction is an important source of risk when driving on two-lane rural roads, the longer the sight distance of the opposite lane, the stronger the driver's sense of safety.

As for environment information, the landscape had the greatest influence on the drivers' speed. "Treetops" and "windows" were found to be crucial information, which reflected the density of roadside vegetation and residence. Charlton et al. [12] demonstrated that roads with abundant greenery and buildings were judged as a destination rather than a direct through-road, and expected speed to be reduced correspondingly. Yu et al. [18] also confirmed that the presence of trees and houses could reduce the probability of speeding. For pavement markings, their location is important, since it takes some time for drivers to react, make decisions, and execute the decisions once they have received such "pulse information".

It was also confirmed that the interaction of environment information cannot be superimposed linearly. The change of one type of information will affect the relative importance of other kinds of information. All categories of information have a significant linear relationship with the type of landscape. Landscape with a slight modification on speed has a high tolerance. In such a case, drivers tend to drive freely, and the relative importance of other information is reduced. As for traffic signs, those with little impact on speed are usually information delivery or danger warning signs, such as "villages ahead", "fast curves ahead", etc. Compared with mandatory signs, these suggestive signs make drivers more aware of the road conditions, so the effectiveness of landscape information is improved.

## 6. Conclusions

This study aimed to quantify the influence of visual road information on driving speed, hoping to inform safer speed choices by optimizing road design. There have been many studies on driving speed. However, drivers' perceptions of the road, as well as engineering practicality, are often ignored. In this study, visual road information was categorized as visual geometry information and visual environment information, which were extracted from RGB images from the driver's perspective. Prediction models for geometry-determined speed and environment-modified speed were established based on the random forest algorithm and a convolutional neural network.

Moreover, the importance of information was also computed. The analysis showed that curvatures of the right boundary in the "middle scene" and the "near scene" were critical factors influencing drivers' speed choice. Landscape can affect driving speed dramatically. It was inferred that there would be two distinct landscapes that could reduce the speed by about 35/h and 5 km/h. Speed changes caused by traffic signs and protection were relatively small. In multi-stimulus scenarios, the effectiveness of these road facilities is often influenced by the category of the landscape.

Compared with previous studies, this study was more concerned with the one-to-one correspondence between specific information and speed. The logic of the study complied with the process of road design; that is, the geometry of the road is selected first, and the landscape and facilities are arranged subsequently. With speed tags as the basis of design, drivers' needs can be better fitted.

One of the limitations of this study was that the models only analyzed driving speeds on rural roads in China; the fitness of the model for other kinds of roads should be further tested. In future study, large-scale field tests could be implemented to obtain more comprehensive driver behavior data. It should be pointed out that the study contributed to road image understanding from the driver's perspective, and proposed an entirely new idea for the design of SER. With the development of intelligent algorithms, it will be possible to realize intelligent optimization of road layout based on driving demand in the future.

**Author Contributions:** Conceptualization, Y.Q. and Y.C.; Data curation, Y.Q. and K.L.; Formal analysis, Y.Q.; Funding acquisition, Y.Q.; Investigation, Y.Q.; Methodology, Y.Q. and K.L.; Project administration, Y.Q.; Resources,

Y.C.; Software, Y.Q. and K.L.; Supervision, Y.C.; Validation, Y.Q.; Visualization, Y.Q.; Writing—original draft, Y.Q.; Writing—review & editing, Y.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Aarts, L.; van Schagen, I. Driving speed and the risk of road crashes: A review. *Accid. Anal. Prev.* **2006**, *38*, 215–224. [[CrossRef](#)] [[PubMed](#)]
2. Tanishita, M.; van Wee, B. Impact of vehicle speeds and changes in mean speeds on per vehicle-kilometer traffic accident rates in Japan. *IATSS Res.* **2017**, *41*, 107–112. [[CrossRef](#)]
3. Charlton, S.G. Conspicuity, memorability, comprehension, and priming in road hazard warning signs. *Accid. Anal. Prev.* **2006**, *38*, 496–506. [[CrossRef](#)] [[PubMed](#)]
4. Zhao, X.; Li, J.; Ding, H.; Zhang, G.; Rong, J. A generic approach for examining the effectiveness of traffic control devices in school zones. *Accid. Anal. Prev.* **2015**, *82*, 134–142. [[CrossRef](#)] [[PubMed](#)]
5. Theeuwes, J.; Godthelp, H. Self-explaining roads. *Saf. Sci.* **1995**, *19*, 217–225. [[CrossRef](#)]
6. Charlton, S.G.; Starkey, N.J. Drivers' mental representations of familiar rural roads. *J. Environ. Psychol.* **2017**, *50*, 1–8. [[CrossRef](#)]
7. Martens, M.H. The failure to respond to changes in the road environment: Does road familiarity play a role? *Transp. Res. Part F Traffic Psychol. Behav.* **2018**, *57*, 23–35. [[CrossRef](#)]
8. Intini, P.; Berloco, N.; Colonna, P.; Ranieri, V.; Ryeng, E. Exploring the relationships between drivers' familiarity and two-lane rural road accidents. A multi-level study. *Accid. Anal. Prev.* **2018**, *111*, 280–296. [[CrossRef](#)]
9. Alonso Plá, F.M.; Esteban Martínez, C.; Calatayud Miñana, C.; Egido Portela, Á. Speed and road accidents: Risk perception, knowledge and attitude towards penalties for speeding. *Psychofenia* **2015**, *31*, 63–76.
10. Weller, G.; Schlag, B.; Friedel, T.; Rammin, C. Behaviourally relevant road categorisation: A step towards self-explaining rural roads. *Accid. Anal. Prev.* **2008**, *40*, 1581–1588. [[CrossRef](#)]
11. Wang, F.; Chen, Y.; Guo, J.; Yu, C.; Stevenson, M.; Zhao, H. Middle-aged drivers' subjective categorization for combined alignments on mountainous freeways and their speed choices. *Accid. Anal. Prev.* **2019**, *127*, 80–86. [[CrossRef](#)] [[PubMed](#)]
12. Charlton, S.G.; Starkey, N.J. Driving on urban roads: How we come to expect the 'correct' speed. *Accid. Anal. Prev.* **2017**, *108*, 251–260. [[CrossRef](#)] [[PubMed](#)]
13. Charlton, S.G.; Mackie, H.W.; Baas, P.H.; Hay, K.; Menezes, M.; Dixon, C. Using endemic road features to create self-explaining roads and reduce vehicle speeds. *Accid. Anal. Prev.* **2010**, *42*, 1989–1998. [[CrossRef](#)] [[PubMed](#)]
14. Mackie, H.; Macmillan, A.; Witten, K.; Baas, P.; Field, A.; Smith, M.; Hosking, J.; King, K.; Sosene, L.; Woodward, A. Te Ara Mua - Future Streets suburban street retrofit: A researcher-community-government co-design process and intervention outcomes. *J. Transp. Health* **2018**, *11*, 209–220. [[CrossRef](#)]
15. Gilandeh, S.S.; Hosseinlou, M.H.; Anarkooli, A.J. Examining bus driver behavior as a function of roadway features under daytime and nighttime lighting conditions: Driving simulator study. *Saf. Sci.* **2018**, *110*, 142–151. [[CrossRef](#)]
16. Antonson, H.; Mårdh, S.; Wiklund, M.; Blomqvist, G. Effect of surrounding landscape on driving behaviour: A driving simulator study. *J. Environ. Psychol.* **2009**, *29*, 493–502. [[CrossRef](#)]
17. Li, Y.; Chen, Y. Driver vision based perception-response time prediction and assistance model on mountain highway curve. *Int. J. Environ. Res. Public Health* **2017**, *14*, 31. [[CrossRef](#)]
18. Yu, B.; Chen, Y.; Bao, S. Quantifying visual road environment to establish a speeding prediction model: An examination using naturalistic driving data. *Accid. Anal. Prev.* **2019**, *129*, 289–298. [[CrossRef](#)]
19. Andueza, P.J. Mathematical models of vehicular speed on mountain roads. *Transp. Res. Rec.* **2000**, 104–110. [[CrossRef](#)]
20. Hamdar, S.H.; Qin, L.; Talebpour, A. Weather and road geometry impact on longitudinal driving behavior: Exploratory analysis using an empirically supported acceleration modeling framework. *Transp. Res. Part C Emerg. Technol.* **2016**, *67*, 193–213. [[CrossRef](#)]

21. Lee, Y.M.; Chong, S.Y.; Goonting, K.; Sheppard, E. The effect of speed limit credibility on drivers' speed choice. *Transp. Res. Part F Traffic Psychol. Behav.* **2017**, *45*, 43–53. [[CrossRef](#)]
22. Gitelman, V.; Pesahov, F.; Carmel, R.; Bekhor, S. The Identification of Infrastructure Characteristics Influencing Travel Speeds on Single-carriageway Roads to Promote Self-explaining Roads. *Transp. Res. Procedia* **2016**, *14*, 4160–4169. [[CrossRef](#)]
23. Maslow, A.H. A theory of human motivation. *Psychol. Rev.* **1943**, *50*, 370–396. [[CrossRef](#)]
24. Yu, B.; Chen, Y.; Fu, Y. Driving Speed Prediction Method for Low-Graded Highways from Drivers' Visual Perception. *J. Tongji Univ. (Nat. Sci.)* **2017**, *45*, 363–370.
25. Yu, B.; Chen, Y.; Bao, S.; Xu, D. Quantifying drivers' visual perception to analyze accident-prone locations on two-lane mountain highways. *Accid. Anal. Prev.* **2018**, *119*, 122–130. [[CrossRef](#)]
26. Yu, B.; Bao, S.; Chen, Y.; Chen, Y. Using 3D mobile mapping to evaluate intersection design through drivers' visual perception. *IEEE Access* **2019**, *7*, 19222–19231. [[CrossRef](#)]
27. Zhao, H.; Qi, X.; Shen, X.; Shi, J.; Jia, J. ICNet for Real-Time Semantic Segmentation on High-Resolution Images. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 418–434.
28. Tang, J.; Zheng, L.; Han, C.; Yin, W.; Zhang, Y.; Zou, Y.; Huang, H. Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review. *Anal. Methods Accid. Res.* **2020**, *27*, 100123. [[CrossRef](#)]
29. Taghi, M.K.; Moiz, G.; Jason, V.H. An empirical study of learning from imbalanced data using Random Forest. In Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), Patras, Greece, 29–31 October 2007.
30. Gong, H.; Sun, Y.; Shu, X.; Huang, B. Use of random forests regression for predicting IRI of asphalt pavements. *Constr. Build. Mater.* **2018**, *189*, 890–897. [[CrossRef](#)]
31. Yu, B.; Bao, S.; Feng, F.; Sayer, J. Examination and prediction of drivers' reaction when provided with V2I communication-based intersection maneuver strategies. *Transp. Res. Part C Emerg. Technol.* **2019**, *106*, 17–28. [[CrossRef](#)]
32. Yoo, H.J. Deep Convolution Neural Networks in Computer Vision: A Review. *IEIE Trans. Smart Process. Comput.* **2015**, *4*, 35–43. [[CrossRef](#)]
33. Alex, K.; Ilya, S.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105.
34. Khan, R.U.; Zhang, X.; Kumar, R. Analysis of ResNet and GoogleNet models for malware detection. *J. Comput. Virol. Hacking Tech.* **2019**, *15*, 29–37. [[CrossRef](#)]
35. Kumar, A.; Kim, J.; Lyndon, D.; Fulham, M.; Feng, D. An Ensemble of Fine-Tuned Convolutional Neural Networks for Medical Image Classification. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 31–40. [[CrossRef](#)] [[PubMed](#)]
36. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.
37. Wang, F.; Chen, F. Influencing factor analysis for visual information search while driving: Based on modeling of visually interesting regions. In Proceedings of the 11th International Conference of Chinese Transportation Professionals (ICCTP), Nanjing, China, 14–17 August 2011.
38. Fu, Y.; Chen, Y. Extracting Image Parameters of Road Environment and Their Impact on Speed. In Proceedings of the 97th Annual Meeting of the Transportation Research Board, Washington, DC, USA, 7–11 January 2018.
39. Chen, F.; Zhou, Z.H.; Yang, Y.X. On driver's viewpoint distribution characteristics on curved section of the highway in mountainous area. *J. Hefei Univ. Technol.* **2015**, *5*, 594–599. [[CrossRef](#)]

