Self-explaining analysis of facility environments on 2-lane rural roads with an improved lightweight CNN considering drivers' visual perception

ARTICLE IN PRESS

International Journal of Transportation Science and Technology xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

journal homepage: www.elsevier.com/locate/ijtst

Research Paper

Self-explaining analysis of facility environments on 2-lane rural roads with an improved lightweight CNN considering drivers' visual perception

Weixi Ren [a], Bo Yu [a,*], Yuren Chen [a], Shan Bao [b,c], Kun Gao [d], You Kong [e]

[a] Key Laboratory of Road and Traffic Engineering of the Ministry of Education, College of Transportation Engineering, Tongji University, 4800 Cao'an Highway, Shanghai 201804, China
[b] Industrial and Manufacturing Systems Engineering Department, University of Michigan-Dearborn, 4901 Evergreen Rd, Dearborn, MI 48128, USA
[c] University of Michigan Transportation Research Institute, 2901 Baxter Rd, Ann Arbor, MI 48109-2150, USA
[d] Department of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg SE-412 96, Sweden
[e] College of Transport and Communications, Shanghai Maritime University, No.1550, Haigang Avenue, Lin'gang Xincheng, Pudong, Shanghai 201303, China

ARTICLE INFO

Article history:
Received 11 April 2024
Received in revised form 4 July 2024
Accepted 4 August 2024
Available online xxxx

Keywords:
:
Self-explaining analysis
Speeding
Improved lightweight convolutional neural network
Drivers' visual perception characteristics
Road category perception

ABSTRACT

Speeding is one of the primary contributors to rural road crashes. Self-explaining theory offers a solution to reduce speeding, which suggests that well-designed facility environments (i.e., road facilities and surrounding landscapes) can automatically guide drivers to choose appropriate speeds on different road categories. This study proposes an improved lightweight convolutional neural network (LW-CNN) that includes drivers' visual perception characteristics (i.e., depth perception and dynamic vision) to conduct the self-explaining analysis of the facility environment on 2-lane rural roads. Data for this study are gathered through naturalistic driving experiments on 2-lane rural roads across five Chinese provinces. A total of 3,502 visual facility environment images, alongside their corresponding operation speeds and speed limits, are collected. The improved LW-CNN exhibits high accuracy and efficiency in predicting operation speeds with these visual facility environment images, achieving a train loss of 0.05% and a validation loss of 0.15%. The semantics of facility environments affecting operation speeds are further identified by combining this LW-CNN with the Gradient-weighted class activation mapping algorithm and the semantic segmentation network. Then, six typical 2-lane rural road categories perceived by drivers with different operation speeds and speeding probability are summarized using k-means clustering. An objective and comprehensive analysis of each category's semantic composition and depth features is conducted to evaluate their influence on drivers' speeding probability and road category perception. The findings of this study can be directly applied to optimize facility environments from drivers' visual perception to decrease speeding-related crashes.

Peer review under responsibility of Tongji University and Tongji University Press.
* Corresponding author.
E-mail address: boyu@tongji.edu.cn (B. Yu).

http://dx.doi.org/10.1016/j.ijtst.2024.08.002
2046-0430/

Please cite this article as: W. Ren, B. Yu, Y. Chen et al., Self-explaining analysis of facility environments on 2-lane rural roads with an improved lightweight CNN considering drivers' visual perception, International Journal of Transportation Science and Technology, http://dx.doi.org/10.1016/j.ijtst.2024.08.002

1. Introduction

Despite low population density and traffic volumes, crashes on rural roads remain both severe and frequent (Wu et al., 2017). Speeding has been recognized as one of the primary contributors to rural road crashes, particularly those resulting in injuries or fatalities (Radam et al., 2022; Job and Brodie, 2022). Thus, reducing the occurrence of speeding on rural roads is of great importance and draws widespread concern (Yu et al., 2019).

There is increasing acknowledgment that the layout of facility environments (i.e., the road facilities and surrounding landscapes) on rural roads has significant effects on drivers' speed selection (Charlton and Starkey, 2017a). Drivers make expectations towards road categories by perceiving facility environments and then adjust their driving speeds accordingly (Charlton et al., 2010). If the perceived facility environments do not match the actual road category, it cannot play the role of effectively guiding driving speeds, potentially leading to increased speeding probability and a heightened risk of crashes, despite drivers' efforts to prevent them.

The self-explaining theory indicates that the facility environment should automatically elicit operation speeds appropriate for the corresponding road category (Theeuwes and Godthelp, 1995). It focuses on creating driver-adapted facility environments using traffic signs, road markings, landscapes, and other semantics, to ensure clear road category perception (Theeuwes, 2021). Since it came out, the self-explaining theory has rapidly gained popularity and has been applied to rural roads (Ghorbani et al., 2023). For example, a study proposed recognizable road markings to emphasize the category of rural roads, and the effectiveness of these markings was confirmed through the collected operation speeds on these roads (Enzfelder, 2013). However, the current self-explaining road design is limited to specific facilities designated by designers, and its effectiveness still remains entirely uncertain if considerable time and effort are not expended for evaluation (De Brucker and Macharis, 2011). The underlying reason for this predicament might be the limited understanding of how facility environments influence drivers' road category perception on rural roads.

Identifying features of facility environments utilized by drivers for road category perception is the key to self-explaining analysis and is crucial in creating self-explaining roads (Walker et al., 2013). Some studies have analyzed these features from the aspects of psychological feelings. For example, based on participants' ratings of various facility environment images, rural roads were classified into three subjective categories: monotony, comfort, and demand, and some distinctive features that differentiate these categories, such as road surface conditions, were identified (Walker et al., 2008). However, the results of these studies are subjective and abstract, making it challenging to generalize universally applicable engineering principles for self-explaining road design. Furthermore, some studies have focused on analyzing the impact of specific semantics on road category perception. For instance, participants were shown facility environment images with semantics that could potentially affect road category perception highlighted and asked to choose appropriate driving speeds (Theeuwes et al., 2023). Then, the relationship between these semantics and driving speeds was analyzed, revealing that the presence of trees and road markings could be key features influencing drivers' road category perception. However, these studies lack a comprehensive consideration of the layout of facility environments. In addition, the current analysis of drivers' road category perception relies heavily on captured images and questionnaire surveys, which have significant differences with real-road driving scenarios (Mårtensson et al., 2018). Currently, there is a lack of an objective and comprehensive method to analyze facility environment features used by drivers for road category perception based on their actual driving behavior.

Drivers' road category perception is mainly a visual search process that integrates drivers' appearance, depth, and dynamic perception of facility environments (Sistu et al., 2019). Depth perception enables drivers to accurately assess the distance of objects within the facility environment, thus facilitating a comprehensive and three-dimensional perception of the facility environment (Choi et al., 2020). The same facility environment semantics at different visual depths can have varying effects on drivers' road category perception. For instance, the research of Pashkevich et al. (2018) demonstrated that road markings with greater visual depth were more attention-grabbing and exhibited more effective driving behavior inducement at open intersections. Additionally, drivers' dynamic perception also plays a crucial role in their road category perception. For example, drivers' gaze areas were found to reduce as driving speeds increased, and traffic signs outside these areas were difficult to notice (Zhang et al., 2010). Considering drivers' visual perception characteristics is essential for a comprehensive understanding of the features influencing road category perception. However, current self-explaining research only focuses on the appearance of facility environments, neglecting the significance of drivers' depth perception and dynamic visual characteristics.

Convolutional Neural Network (CNN), one of the most advanced deep learning techniques, is very effective in image detection and classification (Saleem et al., 2019). Leveraging human-like visual processing mechanisms, CNN can accurately emulate drivers' visual perception during driving and efficiently analyze facility environment images, making it widely employed in traffic-related fields, including traffic sign recognition, collision warning, and pedestrian detection (George, 2019). Compared to traditional manual data extraction methods, CNN enables swift processing of large-scale facility environment images, automatic extraction of relevant features, and unveiling of potentially overlooked features (Comber et al., 2020). CNN has been applied to the self-explaining analysis of facility environments on rural roads. For example, a study by Qin et al. (2020) utilized CNN to analyze the impact of facility environment semantics on operation speeds and identified key features, such as treetops and windows, which significantly decreased operation speeds on rural roads. However, their study only considered the facility environment as a two-dimensional image, neglecting drivers' depth perception and dynamic visual characteristics. Furthermore, the CNN they employed exhibited a large parameter size, making it challenging to apply in engineering practice. With the advancement of computer technology, CNN has witnessed a series of network structure improvements, such as the expansion of input channels and the introduction of lightweight feature extraction modules, which may offer a potential solution to this problem (Wang and Neumann, 2018).

Given the above, the self-explaining theory provides a promising solution to reduce the probability of speeding, by using facility environments to automatically guide drivers to choose appropriate speeds while driving on different road categories. However, there is currently a lack of an objective, comprehensive, and efficient method to investigate how facility environments influence drivers' road category perception and corresponding speed selection. In addition, existing self-explaining research overlooks drivers' visual perception characteristics. Thus, this study put forward an improved LW-CNN that considers drivers' depth perception and dynamic visual characteristics to conduct the self-explaining analysis of facility environments on 2-lane rural roads. This improved LW-CNN is used to predict operation speeds by treating facility environment images as input. Then, this improved LW-CNN is combined with the Gradient-weighted class activation mapping (Grad-CAM) algorithm and a modified semantic segmentation network, in order to further investigate typical 2-lane rural road categories perceived by drivers and identify key features that affect speeding probability and road category perception. We expect that this study can contribute to reducing speeding-related crashes by optimizing the facility environments.

2. Framework

The framework of this study is proposed in Fig. 1, which briefly illustrates the structure of the improved LW-CNN considering drivers' visual perception characteristics, and its application in self-explaining analysis of the facility environment on 2-lane rural roads.

Firstly, integrating the effective visual field extraction based on drivers' dynamic visual characteristics and the depth estimation based on drivers' depth perception characteristics, an improved LW-CNN is proposed to predict the operation speed of the facility environment. Then, the improved LW-CNN is combined with the Grad-CAM algorithm and a modified semantic segmentation network to identify semantics influencing the operation speeds of various facility environments. The typical 2-lane rural road categories perceived by drivers with different operation speeds and speeding probability are summarized using the k-means clustering algorithm, and a comprehensive analysis of the semantic composition and depth features within each category is carried out to evaluate their influence on drivers' speeding probability and road category perception.

3. Methodology

3.1. Naturalistic driving experiment

Naturalistic driving experiments were conducted on 2-lane rural roads during sunny daytime in five Chinese provinces, including Tibet, Anhui, Shandong, Jiangxi, and Zhejiang, covering a total distance exceeding 50,000 km. These roads

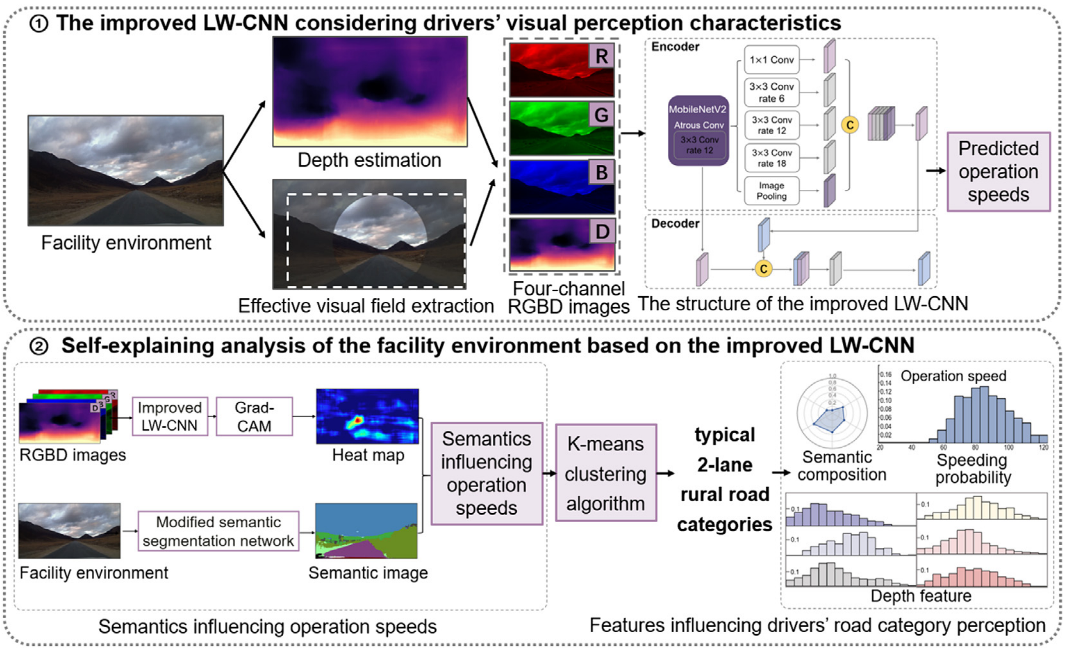


Fig. 1. The improved LW-CNN for self-explaining analysis of the facility environment.

contained diverse road facilities and surrounding landscapes, consisting of mountain roads, township roads, and other categories. A total of 42 experienced drivers (33 males and 9 females) participated in the experiments, with ages ranging from 23 to 50 years (mean = 32.9 years, standard deviation = 7.1 years). All participants had a minimum of three years of driving experience and were familiar with the facility environments. The data collected from each driver included at least driving data of 1,000 km.

The experiments utilized a dual-camera system, comprising GARMIN GDR35 and GARMIN GBC30 driving recorders, to capture the facility environment perceived by drivers. These driving recorders were installed on the vehicle's windshield at the same height as the driver's line of sight. In addition, a GPS locator and a triaxial acceleration sensor were employed to obtain vehicle kinematic information. The data were collected at a sampling rate of 1 Hz. After careful screening to exclude any presence of other road users (including vehicles, non-motor vehicles, pedestrians, and animals), and ensuring adequate illumination and clearly defined speed limits, a total of 3502 valid data sets were obtained. Each data set included stereo image pairs from the drivers' visual point of view, drivers' speeds, and the speed limit.

It should be noted that since this study mainly focuses on the relationship between facility environments and speeding, speeding in this study is defined as operation speeds (i.e., the 85th percentile of all drivers' speeds) exceeding the speed limit (Hou et al., 2020), which indicates that most of the drivers were likely to go overspeed under this facility environment.

3.2. The improved LW-CNN considering drivers' visual perception characteristics

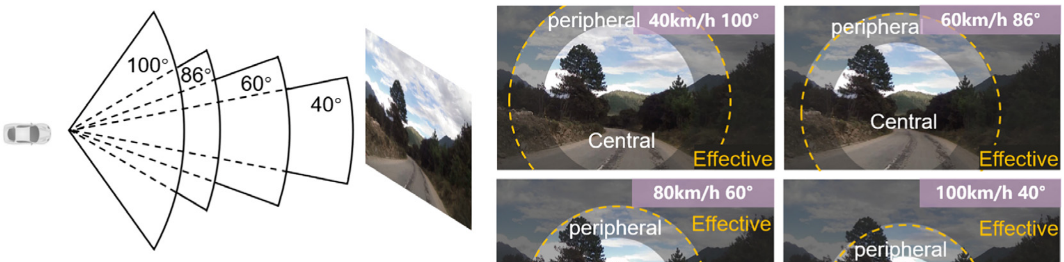
3.2.1. Effective visual field extraction based on drivers' dynamic visual characteristics

During the driving process, not all information in the facility environment can be successfully perceived by drivers. Based on the previous research (Chen and Zheng, 2013), drivers' effective visual field can be divided into two regions: the central visual field and the peripheral visual field. The central visual field represents the area where drivers concentrate their gaze to perceive detailed information about the facility environment, while the peripheral visual field is mainly used to perceive background information about the facility environment. Drivers often neglect information outside the effective visual field. As driving speed increases, drivers tend to focus their gaze on farther areas of the road, resulting in a narrower effective visual field. Fig. 2 illustrates drivers' effective visual field at different driving speeds. For more detailed information about the effective visual field, please refer to our previous study (Chen et al., 2015).

3.2.2. Depth estimation based on drivers' depth perception characteristics

Drivers gain a reliable and accurate perception of the depth of the facility environment by comparing and fusing separated images obtained from left and right eyes (Fujita and Doi, 2016). In this study, considering drivers' depth perception characteristics, PSMNet (i.e., Pyramid Stereo Matching Network) is employed to estimate the depth of the facility environment with a stereo image pair containing left and right images.

PSMNet is a deep learning-based neural network designed for depth estimation (Chang and Chen, 2018). It consists of two weight-sharing pipelines, each comprising a CNN for calculating feature maps, an SPP (Spatial Pyramid Pooling) module for feature extraction, and a convolution layer for feature fusion. The left and right images of the facility environment are processed in parallel to extract their respective features. These features are then utilized to construct a cost volume (height $\times$ width $\times$ disparity $\times$ feature size) for reliable disparity estimation, and depth estimation is performed through regularization and regression on the cost volume using a 3D CNN. Fig. 3 illustrates the network structure of PSMNet. The depth image of facility environments is essentially a grayscale image, where each pixel value represents an estimated depth ranging from 0 to 255. Higher pixel values correspond to greater estimated depths. This study visualizes the grayscale image as an RGB image transitioning from yellow to black for a clearer representation of depths. Closer proximity to yellow indicates a smaller estimated depth, while closer proximity to black indicates a larger estimated depth. Compared to other depth estimation algorithms, PSMNet exhibits advantages such as improved accuracy and shorter runtime (Kokko et al., 2023).



Speed (km/h)	40	60	80	100
Effective visual field (°)	100	86	60	40

Fig. 2. Drivers' effective visual field at different driving speeds.

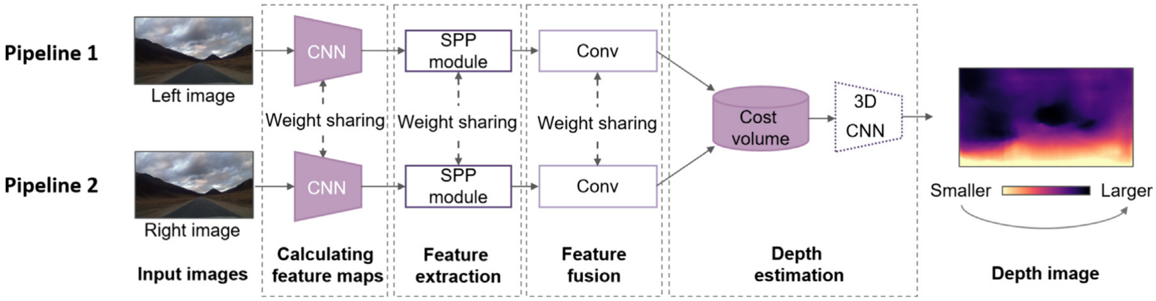


Fig. 3. The network structure of PSMNet.

This study trains PSMNet using the open-source KITTI dataset (Geiger et al., 2013), which includes real stereo image pairs (1,392 $\times$ 512 pixels) and LiDAR point clouds obtained from diverse traffic scenarios through natural driving experiments. To train PSMNet, a subset of 1,504 stereo image pairs is extracted from the KITTI dataset. These image pairs are split into training and validation sets following a ratio of 3:1. The smooth L1 loss function is adopted to train PSMNet, which is commonly used due to its robustness and low sensitivity to outliers (Wang et al., 2020b). PSMNet is implemented using the PyTorch deep learning framework (version 1.13.1). The batch size is set to 32 for training on NVIDIA GeForce RTX 2080TI GPU. The initial learning rate is set to 0.001, considering its widely recognized ability to balance convergence speed and training stability in depth estimation tasks (Chang and Chen, 2018). After training over 500 epochs, stable train and validation losses are observed, indicating sufficient model convergence.

3.2.3. The structure of the improved LW-CNN

This study proposes an improved LW-CNN based on DeepLabV3 + to predict the operation speeds of facility environments on 2-lane rural roads. Fig. 4 presents a comparison between the network structures of the original CNN and the improved LW-CNN. The blue dashed boxes highlight the original network structure, while the red ones emphasize the structure of the improved LW-CNN.

DeepLabV3 + is an advanced CNN proposed by the Google team in 2018 (Chen et al., 2018). As shown in Fig. 4, it adopts an encoder-decoder structure. The encoder is responsible for feature extraction and consists of a deep convolutional neural network with an enhanced Xception network and an Atrous Spatial Pyramid Pooling (ASPP) module. The decoder utilizes upsampling and skip connections to recover the spatial resolution and combine low-level and high-level features for improved prediction accuracy. DeepLabV3 + exhibits several advantages, including high precision, wide applicability, and open-source availability, and it surpasses the most advanced machine learning algorithms in spatial information extraction tasks (Shen et al., 2020). However, DeepLabV3 + still has some limitations. It possesses a large number of parameters that requires significant computational resources and time during both training and inference processes, and its architecture lacks the incorporation of drivers' visual perception characteristics, potentially limiting its applicability in understanding the facility environment (Wang et al., 2021).

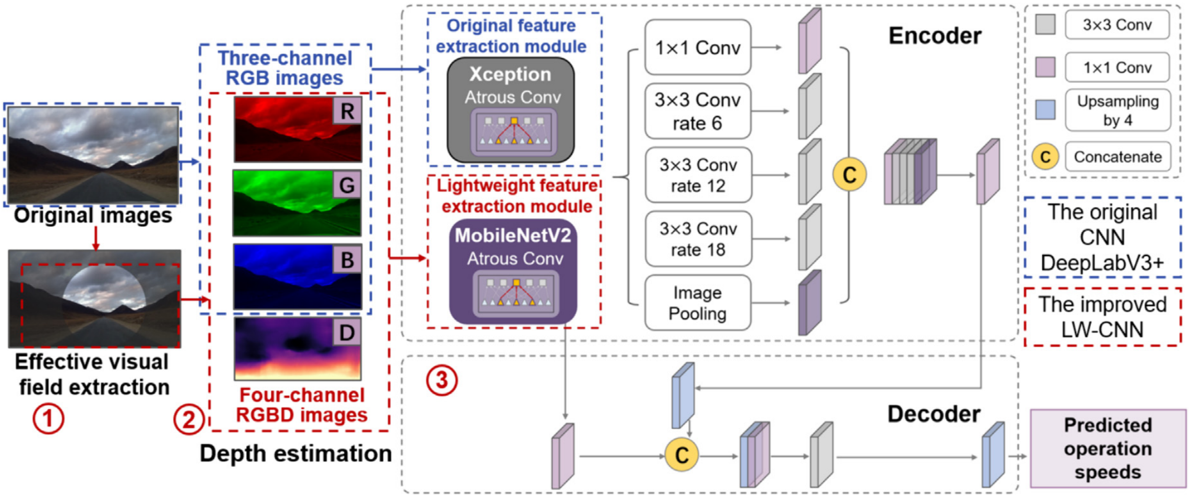


Fig. 4. The structures of the original CNN and the improved LW-CNN.

Therefore, as illustrated in Fig. 4, this study proposes the following three improvements to DeepLabV3+:

(1) The extraction of the effective visual field

To consider drivers' dynamic visual characteristics, an image preprocessing step is introduced into the improved LW-CNN. Specifically, a minimal bounding rectangle, proportionate to the original image, is extracted based on the drivers' effective visual field, and only this region is utilized as the input for the improved LW-CNN.

(2) The consideration of visual depth information

The original DeepLabV3 + model takes RGB images as input, which lacks consideration for visual depth information of facility environments. In this study, RGB images are extended to be four-channel RGBD images with depth images. By incorporating drivers' visual depth information as an additional channel in the input image, the improved LW-CNN could better simulate humans' visual information processing.

(3) The introduction of a lightweight feature extraction module

The feature extraction module of the original DeepLabV3 + model (i.e., Xception) is replaced with MobileNetV2. Compared to Xception, MobileNetV2 has approximately one-sixth the number of parameters and about one-ninety-third of the computational complexity (Chen et al., 2022). The introduction of MobileNetV2 facilitates the reduction of network computational complexity, improves real-time computational efficiency, and expands the model's applicability in practical engineering scenarios (Harkat et al., 2021).

In this study, the improved LW-CNN is used to predict the operation speeds of facility environments based on data collected from natural driving experiments. Drivers' speed perception is not completely precise and is typically represented by speed intervals (Zheng et al., 2018). Consistent with prior research, this study divides the operation speeds into 5 km/h intervals (Sun et al., 2015). These data are split into training and validation sets with a ratio of 3:1. The model is developed using the PyTorch deep learning framework (version 1.13.1). The categorical cross entropy (CCE) is chosen as the loss function. Adam is selected as the network optimizer, with an initial learning rate of 0.001. The training process consists of 300 epochs, and each batch size is set to 4. For evaluating the model's performance, mean accuracy (Mean Acc) is used as the evaluation metric.

3.3. Self-explaining analysis of the facility environment based on the improved LW-CNN

3.3.1. Identification of semantics influencing operation speeds

As shown in Fig. 5, the improved LW-CNN is combined with Grad-CAM and a modified semantic segmentation network to identify facility environment semantics that have influences on operation speeds.

Grad-CAM is a technique proposed to understand and visualize the regions of an input image that contribute significantly to CNN predictions (Selvaraju et al., 2017). By employing global average pooling, the technique assigns weights to features in the rectified convolutional feature maps, highlighting the most influential regions in the decision-making process. These

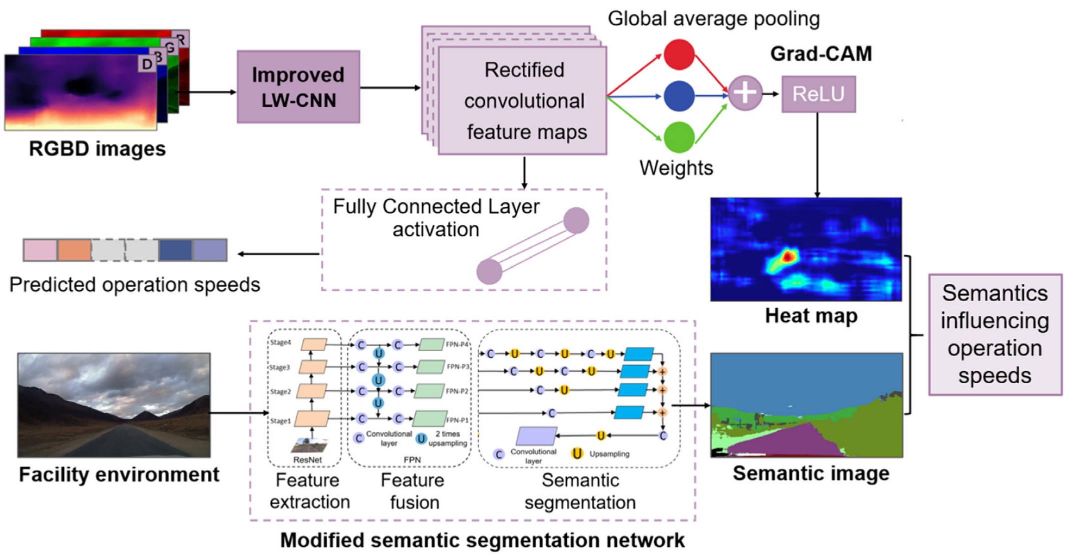


Fig. 5. Identification of semantics influencing operation speeds.

important regions are visualized as heat maps, computed through the rectified linear unit (ReLU). A modified semantic segmentation network is then utilized to semantically annotate these important regions. This network consists of three components: feature extraction, feature fusion, and semantic segmentation. The object features from the facility environment image are retained with the feature extraction module and these features are input into the feature pyramid network (FPN) for feature fusion. Then, the semantic segmentation module identifies various semantics of the facility environment. For further details on the semantic segmentation network, refer to our previous study (Chen et al., 2021).

Based on the above analysis, semantics influencing the operation speeds of the facility environment can be identified. These semantics are classified into six categories: road surfaces, road markings (e.g., dividing lines, edge lines, etc.), traffic signs, roadside protections, natural landscapes, and roadside buildings (e.g., residential areas, schools, etc.). The semantic composition that influences the operation speed of each facility environment can be denoted by a vector $[b_{sur}, b_{mar}, b_{sign}, b_{pro}, b_{lscp}, b_{buid}]$. Elements in the vector describe the six semantic categories' presence, each of which takes a value of 0 or 1. If a semantic category is identified as influencing operation speeds, the corresponding element is marked as 1; otherwise, it is set to 0.

3.3.2. Analysis of drivers' road category perception and their influencing features

The road categories perceived by drivers are analyzed using the k-means clustering algorithm. All facility environments are classified based on their vectors $[b_{sur}, b_{mar}, b_{sign}, b_{pro}, b_{lscp}, b_{buid}]$ to identify typical 2-lane rural road categories. K-means clustering is a commonly used heuristic unsupervised machine learning algorithm for partitioning data sets, which is clear, simple, and efficient (Förster et al., 2019). The Calinski-Harabasz index and the Silhouette coefficient are used to evaluate the clustering performance. The Scikit-learn package in Python is used to perform and test K-means clustering and obtain the class labels for each sample.

Then, this study employs statistical analysis to conduct further self-explaining analysis of facility environments across different typical categories, focusing on analyzing the impact of semantic compositions and depth features on drivers' speeding probability and road category perception.

4. Results

4.1. Performance of the improved LW-CNN

The depth estimation results based on PSMNet are presented in Fig. 6. After 500 training epochs, both the train loss and validation loss consistently decrease. The train loss reaches a low value of 0.35% (See Fig. 6(a)), while the validation loss reduces to 4.05% (See Fig. 6(b)). These results demonstrate the model performs well in depth estimation.

Fig. 7 illustrates the evolution of the loss function of the improved LW-CNN during training for operation speed prediction. The final model achieves a train loss of 0.05% and a validation loss of 0.15%, indicating its good fitting performance and low prediction error.

To evaluate the effectiveness of the proposed improvements on CNN performance in predicting operation speeds, this study compares the results of CNNs with different structures based on Mean Acc and parameter counts, as shown in Fig. 8. These models are structured as follows: Model 1 utilizes MobileNetV2 for feature extraction, with a four-channel RGBD input and effective visual field extraction; Model 2 employs MobileNetV2 for feature extraction, with a three-channel RGB input and effective visual field extraction; Model 3 uses Xception for feature extraction, with a four-channel RGBD input and effective visual field extraction; Model 4 utilizes MobileNetV2 for feature extraction, with a four-channel RGBD input. All these models are trained and tested under the same experimental conditions.

The results demonstrate that incorporating the depth channel leads to an improved Mean Acc of the network. When using the MobileNetV2 feature extraction module with effective visual field extraction, the transition from RGB (Model 2) to RGBD (Model 1) input images results in an increase in Mean Acc from 90.82% to 92.43%, representing a gain of 1.61%.

In addition, the network that replaces the feature extraction module from Xception to MobileNetV2 demonstrates a significant reduction in parameter counts, while experiencing only a slight decrease in Mean Acc. With RGBD input images and

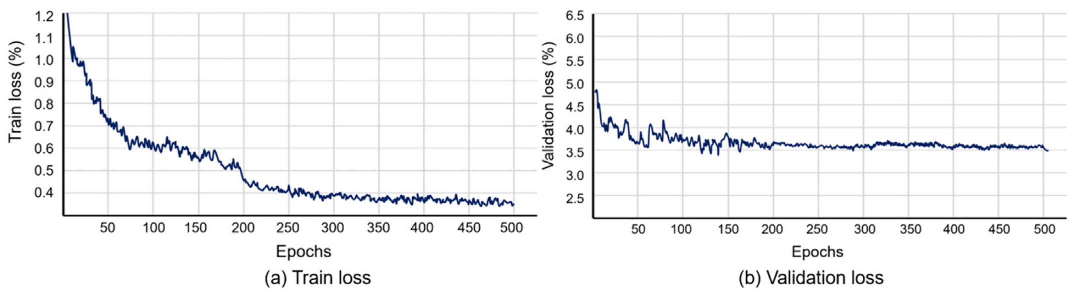


Fig. 6. Evolution of loss function in PSMNet.

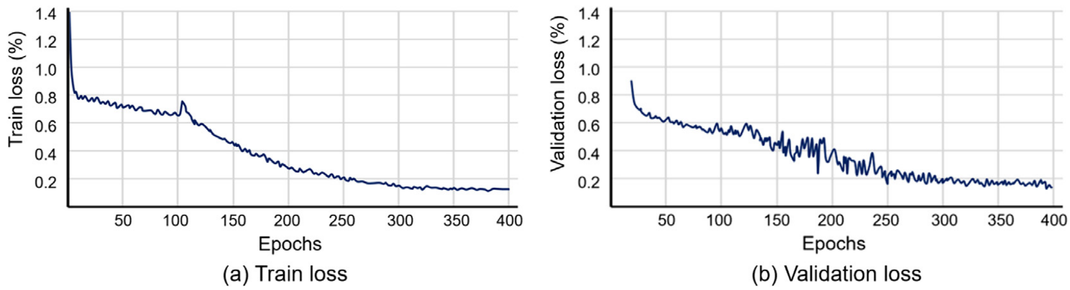


Fig. 7. Evolution of loss function of the improved LW-CNN.

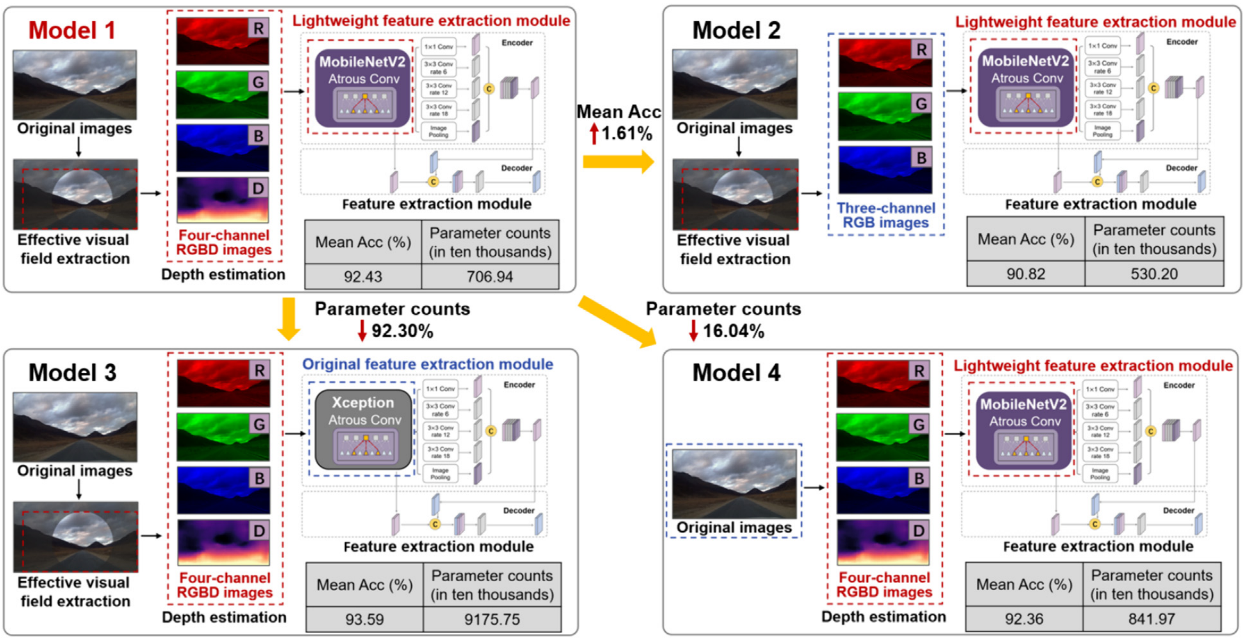


Fig. 8. Performance comparison of CNN Models.

effective visual field extraction, compared to the Xception-based model (Model 3), the MobileNetV2-based model (Model 1) exhibits a reduction in parameter counts by 92.3%.

Besides that, extracting the effective visual field can also significantly reduce the parameter counts, despite having limited impact on the model's Mean Acc. When using RGBD input images with the MobileNetV2 feature extraction module, the model with effective visual field extraction (Model 2) reduces the parameter counts by 16.04% compared to the model without effective visual field extraction (Model 4). Despite this reduction in parameters, both models exhibit almost identical Mean Acc.

Therefore, the improved LW-CNN, employing the MobileNetV2 feature extraction module with RGBD four-channel input images and effective visual field extraction (Model 1), achieves high prediction accuracy while maintaining a lightweight design, making it the most effective model for predicting facility environment operation speeds.

4.2. Self-explaining analysis of the facility environment on 2-lane rural roads

4.2.1. Typical 2-lane rural road categories

In this study, all facility environments are classified with the k-means clustering algorithm to identify typical 2-lane rural road categories. The number of clusters (K) is set from 2 to 10, and the Calinski-Harabasz indexes and the silhouette coefficients under different K values are calculated. When 2-lane rural roads are divided into six categories (i.e., K = 6), the k-means clustering exhibits the best performance with the highest Calinski-Harabasz index and a silhouette coefficient closest to 1, as shown in Fig. 9.

Fig. 10 illustrates the six typical 2-lane rural road categories, each encompassing the following components: 1) An example image of the facility environment in this category; 2) A heat map, which shows regions affecting operation speeds of the facility environment; 3) A depth image, which illustrates the depth distribution of the facility environment; 4) Semantic composition, which summarizes occurrence probability of semantics influencing drivers' road category perception in this category in the form of a radar map; 5) Operation speed, depicting the operation speed probability distribution and the probability of exceeding speed limits (referred as speeding probability, SP) in this category.

These categories are arranged in descending order of SP, denoted as Category 1 to Category 6, and each is named based on its semantic composition features:

(1) Category 1 ("Monotony"): This category includes only the road surface and a few natural surroundings. Drivers exhibit high driving speeds, with a mean speed of 82.60 km/h and an SP of 82%.
(2) Category 2 ("Road marking guidance"): This category mainly comprises road markings, beside the road surface. Drivers tend to travel at a mean speed of 76.5 km/h, with an SP of 66%.
(3) Category 3 ("Multi-facility: traffic sign and road marking"): In this category, in addition to the road surface, there are traffic signs and road markings. Drivers in this category have an SP of 52%, and their mean speed is 67.36 km/h.
(4) Category 4 ("Multi-facility: road marking and roadside protection"): Beside the road surface, this category includes roadside protections and road markings. Drivers exhibit a mean speed of 66.52 km/h, accompanied by an SP of 39%.
(5) Category 5 ("Rich roadside surrounding"): This category includes diverse roadside surroundings, including roadside buildings and natural landscapes. Drivers in this category have an SP of 28%, and the mean speed is 62.63 km/h.
(6) Category 6 ("Multi-facility: traffic sign, roadside protection, and road marking"): This category encompasses various facilities, including traffic signs, roadside protections, and road markings. Drivers exhibit the lowest SP (15%) and have the lowest mean speed of 56.11 km/h.

To verify the significance of road category perception on operation speeds, the one-way analysis of variance (ANOVA) is conducted on the operation speed samples under the six typical 2-lane rural road categories. ANOVA examines whether the differences among groups exceed what can be attributed to chance or random error alone (Kim, 2017). All samples in each category satisfy the required conditions for ANOVA analysis, with normal distributions and homoscedasticity guaranteed (Levene's Test, $p > 0.05$). The results of ANOVA show that road categories have a significant impact on operation speeds ($F(5, 3, 497) = 45.995$, $p < 0.001$). After that, post hoc multiple comparisons are carried out using Tukey's Honestly-Significant Difference test. The results reveal p-values below 0.005 for all pairwise comparisons, indicating a significant distinction in the impacts of different typical 2-lane rural road categories on operation speeds ($p < 0.005$). Then, this study conducts a further analysis of the semantic composition and depth features within each typical 2-lane rural road category, to evaluate their influence on drivers' road category perception.

4.2.2. Semantic composition analysis of different typical 2-lane rural road categories

Different typical 2-lane rural road categories have obvious variations in semantic composition and operation speeds. In all typical 2-lane rural road categories (See Fig. 10(a) to (f)), road surface plays an important role in influencing operation speeds. Existing research also confirms the significant impact of road surface characteristics such as road width or radius, on drivers' road category perception (Ambros et al., 2017). However, due to the limited surrounding land use and capital investment capacity, the space for the improvement of road surfaces on 2-lane rural roads is usually constrained (Coakley et al., 2016). Therefore, this study pays more attention to analyzing the effects of other semantics on drivers' road category perception. From Fig. 10 (a) to (f), it can be observed that as the impact of road surfaces on operation speeds decreases and the influence of other semantics on operation speeds increases, the SP decreases. It suggests that road facilities and surrounding landscapes contribute significantly to drivers' more accurate perception of road categories.

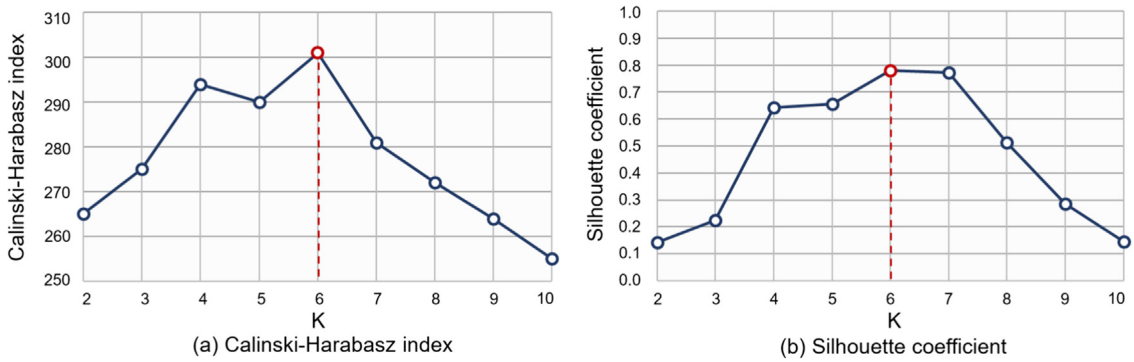


Fig. 9. Results of the k-means clustering algorithm.

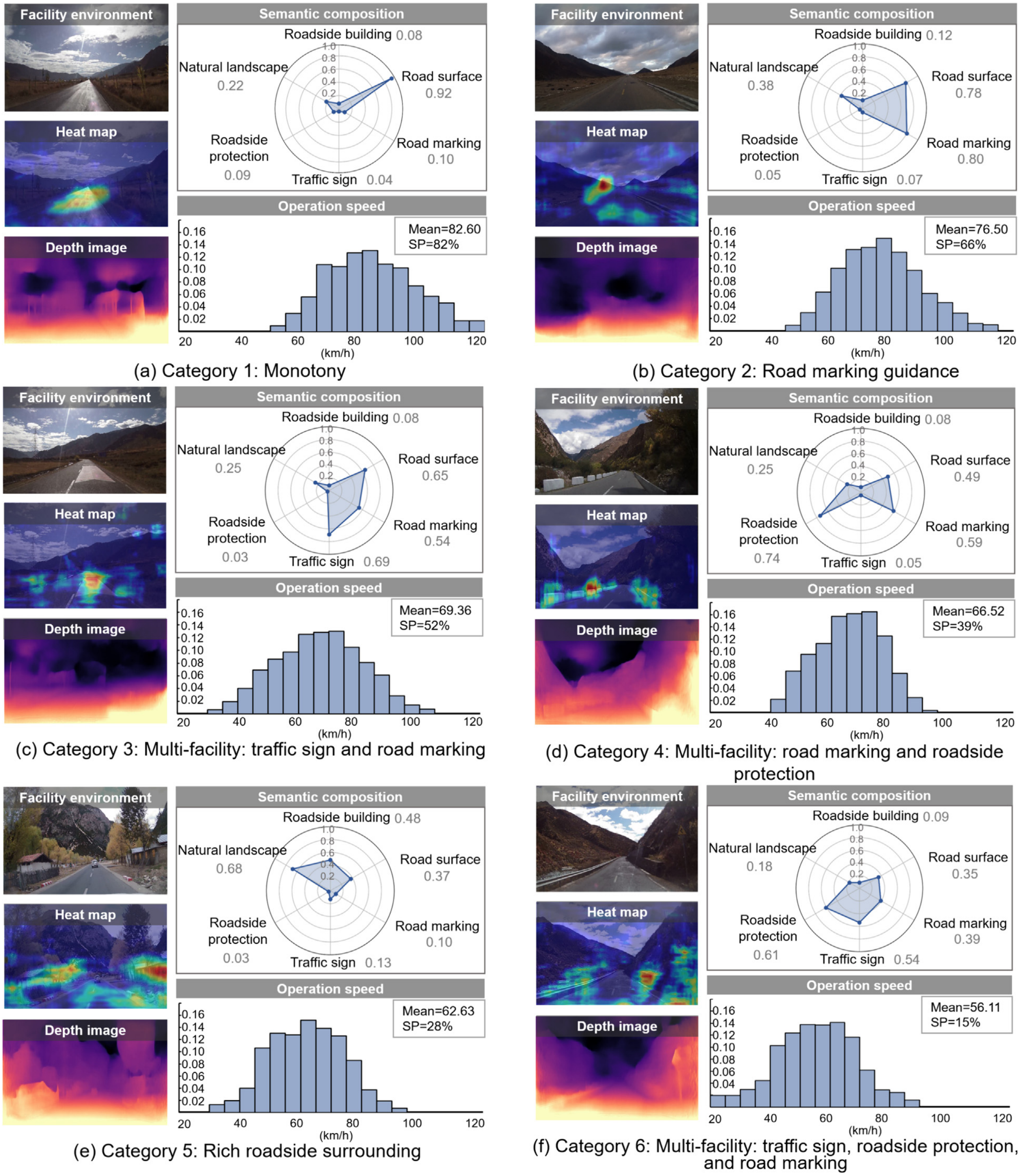


Fig. 10. Six typical 2-lane rural road categories.

In monotonous roads (Category 1, see Fig. 10(a)), drivers often display significantly higher driving speeds, which are not appropriate considering the low-quality conditions of 2-lane rural roads. Monotonous facility environments can decrease driver attentiveness, resulting in severe speeding, which limits drivers' ability to effectively respond to sudden risks on 2-lane rural roads (Ma et al., 2018).

Road markings are the most common road facilities on 2-lane rural roads, providing visual guidance to drivers with lines and symbols. Compared to monotonous roads (Category 1, see Fig. 10(a)), the presence of road markings (Category 2, see

Fig. 10(b)) leads to a decrease of 6.1 km/h in the mean operation speed and a reduction of 16% in the SP. However, the SP remains high (66%) in such facility environments. In comparison to three-dimensional road facilities, road markings are typically distributed along the road and continuously present, which may not sufficiently evoke drivers' risk awareness (Babić et al., 2022). Additionally, non-standard or unclear markings on 2-lane rural roads can further diminish their effectiveness in speed control (Basati and Saeidijam, 2017). Therefore, it is crucial to ensure the standardized design and timely maintenance of these road markings. Additionally, some innovative and eye-catching road markings have been proven to be more effective compared to traditional ones (Babić and Brijs, 2021).

The combination of various traffic facilities can provide a more effective approach to inducing safe driving behavior. When road markings are combined with traffic signs (See Category 3, see Fig. 10(c)), the mean speed and SP decrease by 7.14 km/h and 14%, respectively, compared to Category 2 (See Fig. 10(b)). Traffic signs are symbolic information that delivers precise instructions to drivers through shapes and text (Shinar and Vogelzang, 2013). The combination of traffic signs and road markings balances visual guidance and precise instructions, allowing drivers to form unique road category perceptions. However, their effectiveness appears limited on 2-lane rural roads, since drivers' SP is still over 50% in Category 3.

Compared with traffic signs (Category 3, see Fig. 10(c)), when road markings are combined with roadside protections (Category 4, see Fig. 10(d)), there is a decrease of 2.84 km/h and 13% in the mean speed and SP, respectively. Roadside protections serve not only as physical barriers to prevent vehicles from hazardous areas but also convey iconic meanings to drivers (Bella, 2013). Roadside protections quickly trigger drivers' risk awareness and convey iconic information related to potential hazards, such as sharp curves, steep slopes, and cliffs. Many self-explaining studies have supported the effectiveness of iconic information in creating self-explaining roads and advocated its use over symbolic information on 2-lane rural roads (Theeuwes, 2021).

Rich roadside surroundings (Category 5, see Fig. 10(e)) can significantly influence drivers' road category perception and induce safe driving behavior. Surveys show that drivers subconsciously perceive roads surrounded by abundant roadside buildings and natural landscapes as low-grade and reduce their driving speeds accordingly (Charlton and Starkey, 2017b). Enriching roadside surroundings is a promising self-explaining road design approach, which can not only enhance the safety performance of 2-lane rural roads but also foster the integration of transportation and tourism, thereby promoting rural development (Shukurov et al., 2021).

When road markings, traffic signs, and roadside protections are combined (Category 6, see Fig. 10(f)), the SP and mean speed are the lowest among all typical 2-lane rural road categories. The presence of multiple road facilities provides redundant and complementary visual cues, making it easier for drivers to process and comprehend information related to road conditions and potential hazards (Babić et al., 2020). Furthermore, the combination of road facilities promotes an orderly facility environment, motivating drivers to comply with traffic regulations (Friday, 2012).

4.2.3. Depth feature analysis of different typical 2-lane rural road categories

This study further analyzes the depth features of various semantics that influence drivers' road category perception on 2-lane rural roads. Fig. 11 shows the depth distribution of each semantic in six typical 2-lane rural road categories. The depth of each semantic is categorized based on pixel values in the grayscale image. Specifically, the range of 0 to 85 corresponds to the "near scene", 85 to 170 represents the "mid scene", and 170 to 255 denotes the "far scene".

Categories 1 to 6 demonstrate that the depth of road surfaces has a significant influence on drivers' road category perception, with smaller depths leading to lower mean speeds and reduced SP. When drivers focus their vision on distant road surfaces, the surrounding road details become blurred, resulting in decreased perception of the actual road category. Drivers' focus on road surfaces is largely determined by the facility environment. Research indicates that drivers tend to focus their

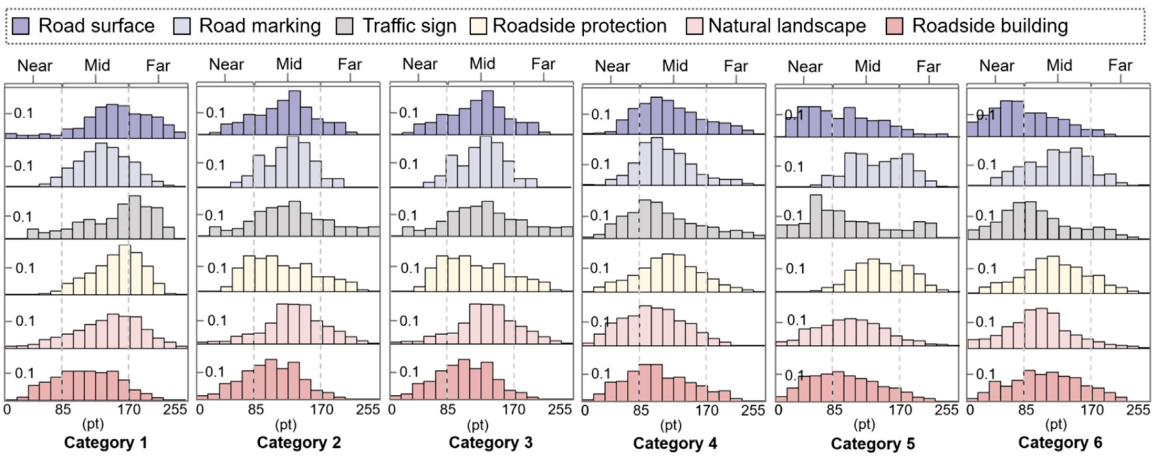


Fig. 11. Depth features of six typical 2-lane rural road categories.

vision on distant road surfaces when the information content of the facility environment is limited (Ma et al., 2020). However, the presence of road facilities and roadside landscapes can enhance the information content of the facility environment, leading drivers to pay more attention to nearby road surfaces and improving their road category perception ability.

Road lanes and roadside protections that influence drivers' road category perception are primarily concentrated in the mid scenes of facility environments. Objects within this scene are prominently positioned in the driver's field of view, making them more effective in visually guiding drivers (Zhou et al., 2023).

Traffic signs influencing drivers' road category perception can be found in near, mid, and far scenes. Compared to other road facilities, traffic signs are elevated above the road surface and are more visible (Doman et al., 2014). It is noteworthy that the depth distributions of traffic signs and road surfaces exhibit some similarity. Consistent with Costa et al.'s findings (2022), the perception depth of noticeable traffic signs is closely associated with the shapes of road surfaces. Therefore, the placement of traffic signs should consider the road surface characteristics to effectively capture drivers' attention.

Roadside buildings and natural landscapes exhibit similar depth features. In Categories 4, 5, and 6, drivers have a relatively good perception of road categories and low SP, and these semantics are close to drivers and mainly found in the near and mid scenes. However, in Categories 1, 2, and 3, drivers have relatively poor perception of road categories and high SP, these semantics are mainly distributed in the mid and far scenes. Research indicates that distant roadside surroundings create an open environment, promoting higher operation speeds (Antonson et al., 2014). Conversely, close roadside surroundings create a sense of an enclosed environment, leading drivers to perceive the road as narrow and limited, resulting in lower operation speeds.

5. Discussion

An increasing number of scholars acknowledge that the "human-oriented" road facility environment design, which meets drivers' expectations, can effectively reduce drivers' tendencies to speeding and consequently lead to an improvement in road safety (Ren et al., 2024). Traditional road facility environment design typically follows a two-dimensional approach, wherein the intricate three-dimensional spatial characteristics of facility environments are primarily decomposed into geometric shapes, longitudinal sections, and cross-sections (Gao et al., 2022). This design method mostly emphasizes vehicle factors, such as design speed and vehicle stability, while often overlooking human factors, especially the visual perception characteristics of drivers (He et al., 2023; Li et al., 2023). However, it is crucial to recognize that most of the information drivers obtain from their surroundings relies on visual perception (Yu et al., 2018). Ignoring drivers' visual perception characteristics may impact the ability of facility environments to enhance drivers' safety awareness and guide driving speeds (Yu et al., 2019). Road design based on drivers' visual perception integrates the traditional two-dimensional design with the three-dimensional road facility environments, ensuring the layout of roads aligns with drivers' visual perception characteristics and behavioral responses, facilitating drivers' accurate road category perception (Wang et al., 2020a; Gao et al., 2024). This study reveals six typical 2-lane rural road categories perceived by drivers and summarizes the impacts of diverse semantic compositions and depth features on drivers' road category perception, operation speeds, and speeding probability. These findings can directly guide road design from drivers' visual perception, contributing to a reduction in speeding probability and an enhancement of 2-lane rural road safety.

With the continuous advancement of computer technology, intelligent techniques have found widespread applications in the evaluation and optimization of road facility environments, such as road surface damage detection using image processing techniques and trajectory data-based identification of crash-prone locations (Tripodi et al., 2020). The complete automation of the road design process has become an inevitable trend (Tang et al., 2020). However, evaluating and optimizing facility environments from drivers' visual perception still demands significant human and material resources and remains relatively inefficient (Vignali et al., 2021). The improved LW-CNN proposed in this study, characterized by its lightweight nature and complemented by Grad-CAM and semantic segmentation techniques, demonstrates the ability to swiftly process and automatically identify critical features within facility environment images. When deployed on mobile or embedded devices, it enables cost-effective automated road risk assessment for 2-lane rural road facility environments from drivers' visual perception.

There are some limitations to this study. The current study only applies to 2-lane rural roads. With the expansion of the dataset, the improved LW-CNN and the self-explaining analysis will be expanded to urban roads, expressways, and various other scenarios. Besides, the improved LW-CNN in this study only considers drivers' depth perception and dynamic visual characteristics. As computer technology and the understanding of human visual perception theories advance, more visual perception characteristics will be integrated into the improved LW-CNN to enhance its environment understanding and feature extraction capabilities.

6. Conclusion

In order to reduce the likelihood of speeding on 2-lane rural roads, this study proposes an improved LW-CNN that considers drivers' visual perception characteristics and applies it to analyze drivers' road category perception and its influencing factors. By integrating important visual perception characteristics, including drivers' depth perception and dynamic visual characteristics, the improved LW-CNN can rapidly process and automatically identify features from facility environment images, and predict operation speeds accurately. Combined with Grad-CAM and semantic segmentation, semantics influencing the operation speed in each facility environment are extracted and visualized in the form of heat maps. After that, k-means clustering is employed to analyze drivers' road category perception on 2-lane rural roads. This study reveals six typical 2-lane rural road categories, each exhibiting significantly different operation speeds and speeding probability. Then, this study conducts a comprehensive analysis of the semantic composition and depth features within each category to explore their impact on drivers' speeding probability and road category perception.

The analysis of semantic composition reveals the significant contribution of road facilities and surrounding landscapes to drivers' accurate perception of road categories and appropriate speed selection. Particularly noteworthy is the combination of road markings, traffic signs, and roadside protections (Category 6), which demonstrates the lowest SP and operation speed compared to all other facility environment categories. Additionally, rich roadside surroundings (Category 5) also have a notable positive influence on drivers' road category perception. The depth feature analysis indicates that drivers' perception of road categories becomes more accurate as the depth of road surfaces decreases, resulting in a reduced probability of speeding. The presence of road facilities and surrounding landscapes can direct drivers' attention to nearby road surfaces, enhancing their road category perception ability. Semantics at specific depths can significantly influence drivers' road category perception. For example, road lanes and roadside protections in the mid scenes have a stronger impact compared to those in the near and far scenes.

This study could contribute to the reduction of speeding probability by offering qualitative and quantitative guidance for 2-lane rural road design and promote the "self-explaining" function of 2-lane rural roads. In addition, the integration of drivers' visual perception characteristics into the LW-CNN enables rapid processing and automatic feature extraction from large-scale facility environment images aligning with human-like visual information processing, which helps automated road risk assessment and offers opportunities for advancing intelligent driving systems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Weixi Ren: Writing – original draft, Methodology, Conceptualization, Writing – review & editing. **Bo Yu:** Writing – original draft, Methodology, Data curation, Conceptualization, Writing – review & editing. **Yuren Chen:** Visualization, Funding acquisition, Conceptualization. **Shan Bao:** Writing – original draft, Methodology, Data curation, Conceptualization. **Kun Gao:** Methodology, Conceptualization. **You Kong:** Writing – original draft, Methodology.

Acknowledgement

This project was jointly supported by the National Natural Science Foundation of China (52102416), the Natural Science Foundation of Shanghai (22ZR1466000), and the Fundamental Research Funds for the Central Universities (22120240159).

Declaration of new publication

The authors declare that the contents of this article has not been published previously.

References

Ambros, J., Valentová, V., Gogolín, O., Andrášik, R., Kubeček, J., Bíl, M., 2017. Improving the self-explaining performance of Czech national roads. Transp. Res. Rec. 2635 (1), 62–70.
Antonson, H., Ahlström, C., Mårdh, S., Blomqvist, G., Wiklund, M., 2014. Landscape heritage objects' effect on driving: A combined driving simulator and questionnaire study. Accid. Anal. Prev. 62, 168–177.
Babić, D., Brijs, T., 2021. Low-cost road marking measures for increasing safety in horizontal curves: A driving simulator study. Accid. Anal. Prev. 153, 106013.
Babić, D., Babić, D., Cajner, H., Sruk, A., Fiolić, M., 2020. Effect of road markings and traffic signs presence on young driver stress level, eye movement and behaviour in night-time conditions: a driving simulator study. Safety 6 (2), 24.
Babić, D., Babić, D., Fiolic, M., Ferko, M., 2022. Road Markings and Signs in Road Safety. Encyclopedia 2 (4), 1738–1752.
Basati, M., Saeidijam, S., 2017. A study of the effects of road deficiencies on traffic accidents in iran. J. Civil Eng. Struct. 1 (1), 35–51.
Bella, F., 2013. Driver perception of roadside con urations on two-lane rural roads: Effects on speed and lateral placement. Accid. Anal. Prev. 50, 251–262.
Chang, J.R., and Chen, Y.S. 2018. Pyramid stereo matching network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Pp. 5410–5418.
Charlton, S.G., Starkey, N.J., 2017a. Drivers' mental representations of familiar rural roads. J. Environ. Psychol. 50, 1–8.
Charlton, S.G., Starkey, N.J., 2017b. Driving on urban roads: How we come to expect the 'correct'speed. Accid. Anal. Prev. 108, 251–260.
Charlton, S.G., Mackie, H.W., Baas, P.H., Hay, K., Menezes, M., Dixon, C., 2010. Using endemic road features to create self-explaining roads and reduce vehicle speeds. Accid. Anal. Prev. 42 (6), 1989–1998.
Chen, Y., Yu, B., and Chen, X. 2015. Urban Road Traffic Visual Environment Reliability Evaluation Based on Amount of Information Perceived by Drivers' Central Vision. In: 25th World Road CongressWorld Road Association (PIARC).

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV). Pp. 801–818.
Chen, Y., Xie, X., Yu, B., Li, Y., and Lin, K. 2021. Multitarget vehicle tracking and motion state estimation using a novel driving environment perception system of intelligent vehicles. J. Adv. Transport. 2021.
Chen, W., Gao, L., Li, X., Shen, W., 2022. Lightweight convolutional neural network with knowledge distillation for cervical cells classification. Biomed. Signal Process. Control 71, 103177.
Chen, Y., Zheng, S., 2013. Mechanism analysis of vehicles operating characteristic affected by visual environment of underground road. J. Tongji Univ. (Nat. Sci.) 07, 1031–1039.
Choi, K., Byun, G., Kim, A., Kim, Y., 2020. Drivers' visual perception quantification using 3D mobile sensor data for road safety. Sensors 20 (10), 2763.
Coakley, R., Storm, R., Neuman, T., 2016. Relationship between geometric design features and performance. Transp. Res. Rec. 2588 (1), 80–88.
Comber, S., Arribas-Bel, D., Singleton, A., Dolega, L., 2020. Using convolutional autoencoders to extract visual features of leisure and retail environments. Landsc. Urban Plan. 202, 103887.
Costa, A.T., Figueira, A.C., Larocca, A.P.C., 2022. An eye-tracking study of the effects of dimensions of speed limit traffic signs on a mountain highway on driverś perception. Transport. Res. F: Traffic Psychol. Behav. 87, 42–53.
De Brucker, K., Macharis, C., 2011. Best things first. The application of multi-criteria analysis to derive implementation priorities for innovative road safety measures. In: Infrastructure and Safety in a Collaborative World: Road Traffic Safety. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 305–325.
Doman, K., Deguchi, D., Takahashi, T., Mekada, Y., Ide, I., Murase, H., Sakai, U., 2014. Estimation of traffic sign visibility considering local and global features in a driving environment. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings. IEEE, pp. 202–207.
Enzfelder, K., 2013. New types of road marking for standardised and self explaining roads. Prace Naukowe Politechniki Warszawskiej. Transport 96, 171–180.
Förster, D., Inderka, R.B., Gauterin, F., 2019. Data-driven identification of characteristic real-driving cycles based on k-means clustering and mixed-integer optimization. IEEE Trans. Veh. Technol. 69 (3), 2398–2410.
Friday, D., 2012. Modification of road infrastructure to increase driver compliance in Uganda. Inter. J. Adv. Manage. Econom. 1 (3), 78–87.
Fujita, I., Doi, T., 2016. Weighted parallel contributions of binocular correlation and match signals to conscious perception of depth. Philos. Trans. R. Soc., B 371, (1697) 20150257.
Gao, Y., Gao, T., Wu, Y., Wang, P., He, Q., 2022. Low-construction-emission cross-section optimization for mountainous highway alignment designs. Transp. Res. Part D: Transp. Environ. 105, 103249.
Gao, J., Yu, B., Chen, Y., Bao, S., Gao, K., Zhang, L., 2024. An ADAS with better driver satisfaction under rear-end near-crash scenarios: A spatio-temporal graph transformer-based prediction framework of evasive behavior and collision risk. Transport. Res. Part C: Emerging Technol. 159, 104491.
Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The kitti dataset. Inter. J. Robot. Res. 32 (11), 1231–1237.
George, C.P., 2019. Convolutional neural networks: alternate drivers' visual perception. IEEE Potent. 39 (1), 19–24.
Ghorbani, M., Saffarzadeh, M., Naderan, A., 2023. Crash prediction modeling for horizontal curves on two-lane, two-way rural highways based on consistency and self-explaining characteristics using zero-truncated data. KSCE J. Civ. Eng., 1–14
Harkat, H., Nascimento, J.M., Bernardino, A., 2021. Fire detection using Deeplabv3+ with mobilenetv2. In: IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, pp. 4095–4098.
He, L., Yu, B., Chen, Y., Bao, S., Gao, K., Kong, Y., 2023. An interpretable prediction model of illegal running into the opposite lane on curve sections of two-lane rural roads from drivers' visual perceptions. Accid. Anal. Prev. 186, 107066.
Hou, W., Lyu, N., Liu, Z., Wang, X., 2020. Modelling large vehicles operating speed characteristics on freeway alignment based on aggregated GPS data. IET Intel. Transport. Syst. 14 (8), 857–865.
Job, R.S., Brodie, C., 2022. Road safety evidence review: Understanding the role of speeding and speed in serious crash trauma: A case study of New Zealand. J. Road Safety 33 (1), 5–25.
Kim, T.K., 2017. Understanding one-way ANOVA using conceptual figures. Korean J. Anesthesiol. 70 (1), 22–26.
Kokko, M.A., Andersen, L.C., Halter, R.J., 2023. Evaluating a PSMNet-based pipeline for intraoperative stereo reconstruction with the da Vinci Xi. Medical Imaging 2023: Image-Guided Procedures, Robotic Interventions, and Modeling, Vol. 12466. SPIE, pp. 71–75.
Li, Z., Yu, B., Wang, Y., Chen, Y., Kong, Y., Xu, Y., 2023. A novel collision warning system based on the visual road environment schema: An examination from vehicle and driver characteristics. Accid. Anal. Prev. 190, 107154.
Ma, J., Gu, J., Jia, H., Yao, Z., Chang, R., 2018. The relationship between drivers' cognitive fatigue and speed variability during monotonous daytime driving. Front. Psychol. 9, 459.
Ma, Y., Qi, S., Zhang, Y., Lian, G., Lu, W., Chan, C.Y., 2020. Drivers' visual attention characteristics under different cognitive workloads: An on-road driving behavior study. Int. J. Environ. Res. Public Health 17 (15), 5366.
Mårtensson, H., Keelan, O., Ahlström, C., 2018. Driver sleepiness classification based on physiological data and driving performance from real road driving. IEEE Trans. Intell. Transp. Syst. 20 (2), 421–430.
Pashkevich, A., Burghardt, T.E., Shubenkova, K., Makarova, I., 2018. Analysis of drivers' eye movements to observe horizontal road markings ahead of intersections. In: Conference Vision Zero for Sustainable Road Safety in Baltic Sea Region. Springer International Publishing, Cham, pp. 1–10.
Qin, Y., Chen, Y., Lin, K., 2020. Quantifying the effects of visual road information on drivers' speed choices to promote self-explaining roads. Int. J. Environ. Res. Public Health 17 (7), 2437.
Radam, I.F., Harianto, D., Narang, R.I., 2022. Factors causing traffic accidents based on the trip maker perception: comparison between urban and rural roads. J. Southwest Jiaotong Univ. 57 (3).
Ren, W., Yu, B., Chen, Y., Gao, K., Bao, S., Wang, Z., Qin, Y., 2024. An intelligent optimization method for the facility environment on rural roads. Comput. Aided Civ. Inf. Eng.
Saleem, M.H., Potgieter, J., Arif, K.M., 2019. Plant disease detection and classification by deep learning. Plants 8 (11), 468.
Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 618–626.
Shen, C., Liu, L., Zhu, L., Kang, J., Wang, N., Shao, L., 2020. High-throughput in situ root image segmentation based on the improved DeepLabv3+ method. Front. Plant Sci. 11, 576791.
Shinar, D., Vogelzang, M., 2013. Comprehension of traffic signs with symbolic versus text displays. Transport. Res. F: Traffic Psychol. Behav. 18, 72–82.
Shukurov, M., Nurdinov, M., Soliev, B., Begboyev, F., 2021. Highways, functions and importance. Am. J. Eng. Technol. 3 (04), 1–6.
Sistu, G., Leang, I., Chennupati, S., Yogamani, S., Hughes, C., Milz, S., Rawashdeh, S., 2019. Neurall: Towards a unified visual perception model for automated driving. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE, pp. 796–803.
Sun, R., Zhuang, X., Wu, C., Zhao, G., Zhang, K., 2015. The estimation of vehicle speed and stopping distance by pedestrians crossing streets in a naturalistic traffic environment. Transport. Res. F: Traffic Psychol. Behav. 30, 97–106.
Tang, F., Ma, T., Zhang, J., Guan, Y., Chen, L., 2020. Integrating three-dimensional road design and pavement structure analysis based on BIM. Autom. Constr. 113, 103152.
Theeuwes, J., 2021. Self-explaining roads: What does visual cognition tell us about designing safer roads? Cognit. Res.: Princ. Implicat. 6 (1), 15.
Theeuwes, J., Snell, J., Koning, T., and Bucker, B. 2023. Self-Explaining Roads: Effects of road design on speed choice.
Theeuwes, J., Godthelp, H., 1995. Self-explaining roads. Safety Sci. 19 (2–3), 217–225.
Tripodi, A., Mazzia, E., Reina, F., Borroni, S., Fagnano, M., Tiberi, P., 2020. A simplified methodology for road safety risk assessment based on automated video image analysis. Transp. Res. Procedia 45, 275–284.

Vignali, V., Acerra, E.M., Lantieri, C., Di Vincenzo, F., Piacentini, G., Pancaldi, S., 2021. Building information Modelling (BIM) application for an existing road infrastructure. Autom. Constr. 128, 103752.

Walker, G.H., Stanton, N.A., Chowdhury, I., 2013. Self Explaining Roads and situation awareness. Saf. Sci. 56, 18–28.

Wang, F., Chen, Y., Wijnands, J.S., Guo, J., 2020a. Modeling and interpreting road geometry from a driver's perspective using variational autoencoders. Comput. Aided Civ. Inf. Eng. 35 (10), 1148–1159.

Wang, C., Du, P., Wu, H., Li, J., Zhao, C., Zhu, H., 2021. A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net. Comput. Electron. Agric. 189, 106373.

Wang, Q., Ma, Y., Zhao, K., Tian, Y., 2020b. A comprehensive survey of loss functions in machine learning. Annals of Data Sci., 1–26

Wang, W., Neumann, U., 2018. Depth-aware cnn for rgb-d segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 135–150.

Weller, G., Schlag, B., Friedel, T., Rammin, C., 2008. Behaviourally relevant road categorisation: A step towards self-explaining rural roads. Accid. Anal. Prev. 40 (4), 1581–1588.

Wu, L., Lord, D., Geedipally, S.R., 2017. Developing crash modification factors for horizontal curves on rural two-lane undivided highways using a cross-sectional study. Transp. Res. Rec. 2636 (1), 53–61.

Yu, B., Chen, Y., Bao, S., Xu, D., 2018. Quantifying drivers' visual perception to analyze accident-prone locations on two-lane mountain highways. Accid. Anal. Prev. 119, 122–130.

Yu, B., Chen, Y., Bao, S., 2019. Quantifying visual road environment to establish a speeding prediction model: an examination using naturalistic driving data. Accid. Anal. Prev. 129, 289–298.

Zhang, C., Yang, S., Pan, B., Zhao, Y., 2010. Safety evaluation of expressway alignment based on spatial valid vision. J. Highway Transp. Res. Develop. 27 (9), 132–137.

Zheng, Z., Du, Z., Xiang, Q., Chen, G., 2018. Influence of multiscale visual information on driver's perceived speed in highway tunnels. Adv. Mech. Eng. 10, (12) 1687814018818961.

Zhou, Y., Tang, L., Huang, J., Xiang, Y., Ge, Y., 2023. How the position distribution of HUD information influences the driver's recognition performance in different scenes. In: International Conference on Human-Computer Interaction. Springer Nature Switzerland, Cham, pp. 383–395.

Further Reading

Dumbaugh, E., Saha, D., Merlin, L., 2020. Toward safe systems: Traffic safety, cognition, and the built environment. J. Plan. Educ. Res. 0739456X20931915