Wang *et al. Heritage Science* (2024) 12:36
https://doi.org/10.1186/s40494-024-01145-0

Heritage Science

RESEARCH

Open Access

Immersive virtual reality and computer vision for heritage: visual evaluation and perception of the industrial heritage sites along the Yunnan–Vietnam railway (Yunnan section)

Yingxue Wang[1], Shuoyi Wang[2], Yue Pan[1], Chunhua Li[1], Chunxu Chen[1] and Jin Wang[1*]

Abstract

The visual composition and human perception are found to relate to the reuse and tourism of heritage railways. Previous studies have used either environmental audits and on-site interviews that have limitations in terms of cost, time, and measurement scale, or virtual perception base on two-dimension images but with gaps in interactivity, virtual immersion and field of view. This study developed an "objective + subjective" visual evaluation and perception framework integrating Computer Vision (CV) and Immersive Virtual Reality (IVR) to assess the visual quality of industrial heritage sites along the Yunnan-Vietnam Railway (Yunnan Section). The stepwise multiple linear regression models were carried out to investigate the relationship between objective evaluation and subjective perception. The results showed that 16 landscape elements of the heritage sites were successfully segmented. According to the visual perception score bands, the 120 industrial heritage sites were classified as 39 high-score sites, 66 medium-score sites, and 15 low-score sites. In general, although the sky and hard ground accounted for a higher proportion, they had little effect on the sum scores, while the vegetation, water, and buildings played a significant role in the perception of visual quality. The results can help researchers, planners, and government departments clarify the visual quality to scientifically specify bottom-up planning and management solutions for railway industrial heritage sites. Moreover, the simplicity, accuracy, and effectiveness of this framework make it suitable for large-scale visual evaluation of other railway industrial heritage sites and linear heritage sites.

Keywords Immersive virtual reality, Computer vision, Visual evaluation and perception, Industrial heritage sites, The Yunnan-Vietnam Railway (Yunnan Section)

*Correspondence:
Jin Wang
jeanwang@swfu.edu.cn
[1] College of Landscape Architecture and Horticulture Sciences, Southwest Forestry University, Kunming 650224, China
[2] School of Art and Product Design, Yunnan Light and Textile Industry Vocational College, Kunming 650300, China

Introduction

Heritage railways serve as tangible remnants of human industrial civilization and have made significant contributions to global transportation, economic progress, and cultural exchange [1]. Presently, the United Nations Educational, Scientific and Cultural Organization (UNESCO) has officially inscribed four heritage railways on the World Heritage list, including the Semmering Railway, the Mountain Railways of India, the Rhaetian Railway in the Albula / Bernina Landscapes, and the

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Trans-Iranian Railway. Since 2018, a total of 14 heritage railways, including the Middle East Railway and Yunnan-Vietnam Railway, have been consecutively enlisted in China's Industrial Heritage Protection List. These railways span across numerous provinces in China, collectively extending over a distance exceeding 9700 km. Some of these railways have been converted into theme parks, museums, greenways, tourist facilities, and green infrastructure through landscape remediation efforts. However, certain transformations lacking scientific analysis have worsened the degradation of this valuable heritage [2]. Additionally, repurposing railway heritage as tourist attractions is a prominent strategy in the tourism industry to generate economic benefits, compete with alternative modes of transportation, and enrich the local cultural identity [3, 4]. Railway tourism is considered to be a significant factor in addressing poverty and promoting sustainable development in developing countries such as China and Vietnam. As a major component of the human five senses for people to perceive the landscape, visual information had an effect of 76% on overall satisfaction [5, 6], which is considered to be an important foundation to support landscape improvement and tourism development on heritage railways [7, 8]. The measurement of objective visual evaluation and subjective visual perception of heritage sites, as well as the examination of their associations, hold significant theoretical and practical implications.

The increasing attention given to landscape elements in relation to landscape visual quality has garnered interest among scholars due to their influential role in shaping individuals' subjective perceptions of landscapes. Additionally, these elements are crucial factors in the realms of tourism and the repurposing of heritage sites [3, 9]. With the increased development and popularity of deep learning technology, scene parsing based on semantic segmentation models has been used to determine the proportion of features that can be seen in a scene [10, 11]. As a core topic in computer vision, semantic segmentation aims to assign a label to each pixel in an image [12]. Many ready-to-use models, such as pyramid scene parsing network (PSPNet), SegFormer, and DeepLab, can be used for image segmentation [13–15]. The assessment of visual quality in a landscape can be conducted by evaluating the proportion of visible elements [16]. This approach has been extensively employed in assessing the visual quality of both natural and cultural environments, such as streets [17, 18], communities [19] and parks [20], and heritage sites [21]. However, there is a dearth of semantic segmentation research specifically focused on railroad industrial heritage sites. Moreover, relying solely on the proportion of image elements for landscape assessments has inherent limitations in terms of reliability and effectiveness, as human subjective visual perception also plays a crucial role.

There has been a long history of research on landscape visual perception [22, 23]. Photo-based surveys and on-site interviews have been widely used to gather stated preferences [24, 25]. Although there has been a proliferation in the use of two-dimensional (2D) media (such as photos) for visual evaluation as an on-field audit alternative [26, 27], there are limitations in terms of interactivity, virtual immersion, and vision. IVR displays enable participants to immerse themselves in a virtual environment by providing a panoramic image centered on the camera [28]. Additionally, the use of iPads and smartphones equipped with gyroscopes and accelerometers allows for remote viewing of virtual environments. Consequently, IVR technologies offer a more intuitive experience for landscape visual perception compared to traditional image-based evaluation methods [29, 30]. Researchers have examined the types of human perception in urban and forest areas, including safety, beauty, color, liveliness, boredom, and depression [31, 32]. A previous IVR-related canal heritage perception study selected beauty, pleasure, tranquility, color, complexity, and liveliness as visual perception indicators [33]. However, the existing visual perception indicators for heritage related to IVR primarily focus on the environment, neglecting the unique characteristics of heritage.

The combination of IVR and CV technologies overcome the limitations of previous methods and show great potential for studying visual quality [34]. A more cost-effective alternative to conducting on-site perception audits is to immerse oneself in a 360-degree view of the heritage site, which has demonstrated comparable feasibility and effectiveness [35, 36]. The semantic segmentation model allows batch analysis of panorama landscape elements, which facilitates the analysis of the relationship between the visual composition and visual perception [10, 37, 38]. Some scholars have endeavored to investigate the assessment of visual landscape quality through the integration of computer vision (CV) and immersive virtual reality (IVR) technologies, with a focus on tourism and heritage preservation. However, the measurement of landscape quality remains challenging due to the subjective preferences of users, especially in intricate environments and regions with extensive cultural and social backgrounds [39]. This study aims to develop an "objective + subjective" visual evaluation and perception framework for industrial heritage sites along the Yunnan-Vietnam Railway (Yunnan Section) by combining CV and IVR technologies. The results can help researchers, practitioners, and government departments clarify the visual quality to scientifically specify bottom-up planning and management solutions for railway industrial heritage sites. The research questions are as follows:

- How can an “objective + subjective” visual evaluation and perception framework be operated for large-scale railway heritage sites based on panoramas?
- What are the differences in their objective visual composition, and what are the differences in people’s subjective perceptions among various railway industrial heritage sites?
- Does objective visual evaluation correlate with subjective visual perception?

Material and methods

Study area

The Yunnan-Vietnam Railway, connecting Haiphong (the largest port city in northern Vietnam) and Kunming (the capital city of Yunnan, China), is the first alpine narrow-gauge railway in China and was nominated as a national industrial heritage site in 2018[3, 40]. The existing foreign research on the Yunnan-Vietnam Railway primarily focuses on historical and cultural analysis, which predominantly relies on foundational information and historical records from France, such as books, drawings, documentaries, and photographs. Various studies conducted in China have explored different aspects of railway heritage, including the economy, history, culture, and the preservation of heritage resources across various periods, regions, and segments of railroads. These studies have also touched upon topics such as tourism development, value assessment, and the construction of heritage corridors for railway heritage. Railway industrial heritage sites serve as significant remnants of heritage railways, encapsulating the origins, growth, and transformation of these historical rail systems. However, there is a noticeable gap in research concerning the visual aspects of railway heritage sites, which consequently hinders the availability of scientific support for certain tourism development initiatives and landscape enhancements. In view of operability, this study selected 120 industrial heritage sites along the Yunnan-Vietnam Railway (Yunnan Section) and Gebishi Railway as the sample by removing those that have been demolished or are inaccessible to the COVID-19 epidemic. The sample includes 78 stations (ID: 1–78), 11 bridges (ID: 79–89), 5 residential heritage sites (ID: 90–94), 15 public service heritage sites (ID: 95–109), and 11 productive heritage sites (ID: 110–120) (Fig. 1, Appendix A).

Research framework

The research framework comprises four stages (Fig. 2). Initially, IVR panoramas were gathered for each sample site. Subsequently, PSPNet model was employed to compute the pixel ratio of each visual element as an objective feature of the eye-level heritage environment. Thirdly, the panorama image was visualized using the PICO Neo3 Head-mounted Display, enabling participants to immerse themselves in the heritage site for subjective visual perception. Finally, the stepwise multiple linear regression models were conducted to examine the relationship between objective evaluation and subjective perception.

Data collection

A Canon 5d3 camera equipped with an EF8- 15 mm f/4L (USM) fisheye lens was used to obtain the panoramas (from August to December 2021). We aimed for consistency in the weather and angles of the shoot as much as possible. We first adjust the shooting angle to 0°and take a picture every 60° when rotating clockwise to obtain a 360° image in the horizontal direction. Then, we adjust the shooting angle to -45° and + 45°respectively, and repeat the above operation to obtain a 180-degree image in the vertical direction. The camera was in line with the sitting height of the human eye (1.2 m) while shooting, and 1to 3 panoramic images were taken at each heritage site. We obtained 205 panoramas (300dpi) after stitching and adjusting by Photomatix Pro 6.2.1, Adobe Photoshop Lightroom, PTGui, Pano2IVR6, and Adobe Photoshop CC2019. However, it may be difficult for participants to view all panoramas. Therefore, a total of 120 panoramas were selected after discussion (1panoramic image for each heritage site).

Objective visual analysis

Semantic segmentation based on the PSPNet model

Semantic segmentation is a fundamental part of computer vision for parsing scenes. Many ready-to-use models, such as PSPNet, SegFormer, and DeepLab, can be used for image segmentation [13–15]. DeepLab and PSPNet are commonly used models for semantic segmentation. After conducting a comprehensive comparison of the segmentation effects between DeepLab and PSPNet models on railway heritage sites using Python 3.7 in PyCharm, the PSPNet model was finally chosen (Fig. 3). The pyramid pool module is the core module of PSPNet, which can aggregate image information of different scales and improve the ability to obtain multiscale features. The workflow of PSPNet was as follows: In Step 1, we gave an image; In Step 2, the CNN was applied to obtain the feature map of the last convolutional; In Step 3, a pyramid parsing module was used to harvest different representations of subregions; In Step 4, the representation was fed into a convolution layer to obtain the final per-pixel prediction [41].

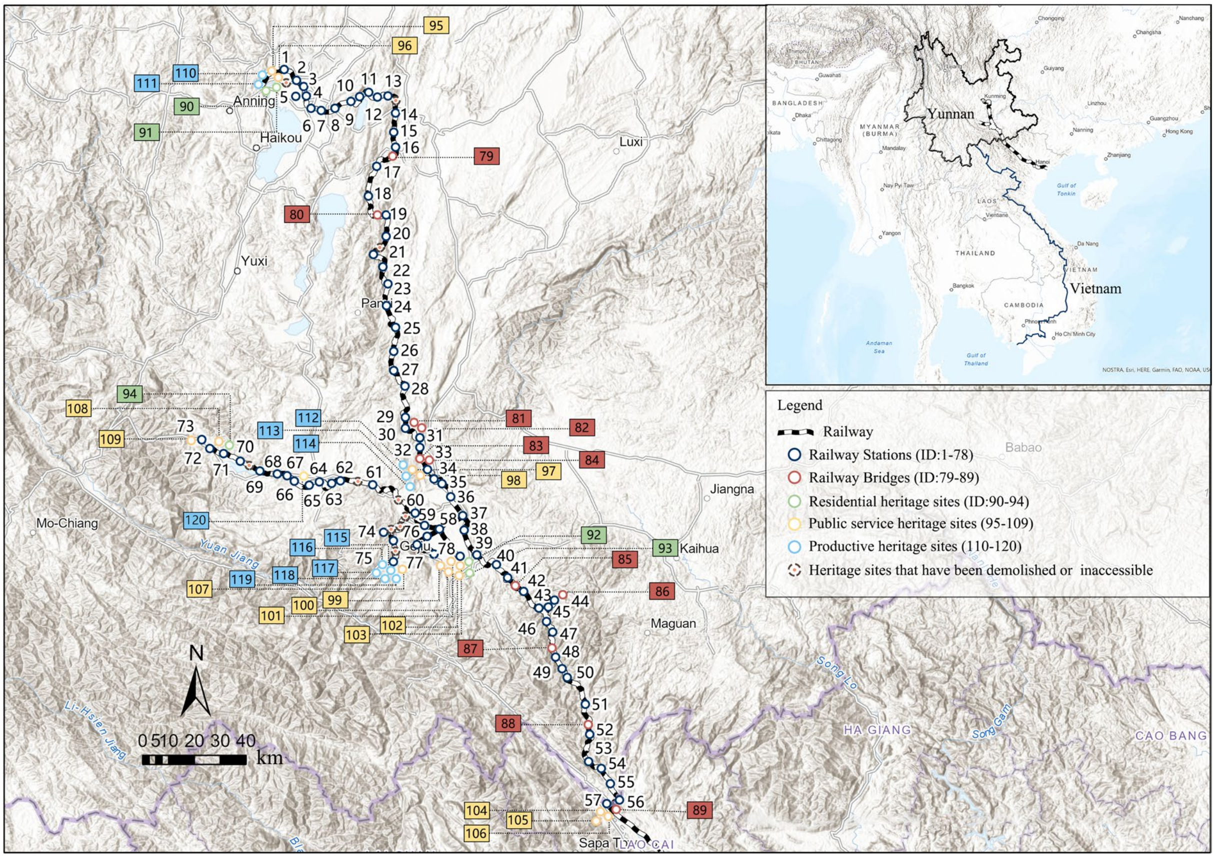

Fig. 1 Study area and location ID

Objective index system

Aesthetic perception and visual quality are profoundly influenced by both natural and artificial elements (such as vegetation, water, and buildings) [20]. Previous studies usually chose one or more indices to merge the element pixels in the image, among which the green visibility index (GVI) is one of the most commonly used indices [17]. Based on the industrial heritage visual evaluation research [33] and combined with railway characteristics, the visibility indices of landscape elements were grouped into six indices: GVI, water visibility index (WVI), sky visibility index (SKVI), hard ground visibility index (HVI), buildings visibility index (BVI), and other elements visibility index (OVI) (Table 1). GVI, WVI, and SKVI are natural indices that can be used to measure the natural landscape ecological environment of the heritage sites. HVI, BVI, and OVI are artificial indices that can reflect the intensity of artificial construction around the heritage sites. The selected indices cover the elements contained in the heritage and can objectively reflect the visual condition of the heritage landscape.

Subjective visual perception

Index system and questionnaire

Previous research has investigated multiple dimensions of human perception, such as beauty, color, and pleasure, in the context of landscape perception in virtual reality. Similarly, studies on linear landscapes and urban parks have employed indices such as color, space, culture, tranquility, and beauty [33, 34]. Adopting the previous experience in the state of the art of visual perception and the characteristics of the railway heritage sites, space, color, texture, uniqueness, culture, history, beauty, and pleasure were selected as visual perception indices. Beauty and pleasure are commonly used indices in landscape visual quality assessment, which can describe the aesthetic preference of the public [42]. Space, color, and texture are mainly used to perceive the environmental texture, while uniqueness, culture, and history are mainly applied to perceive heritage characteristics.

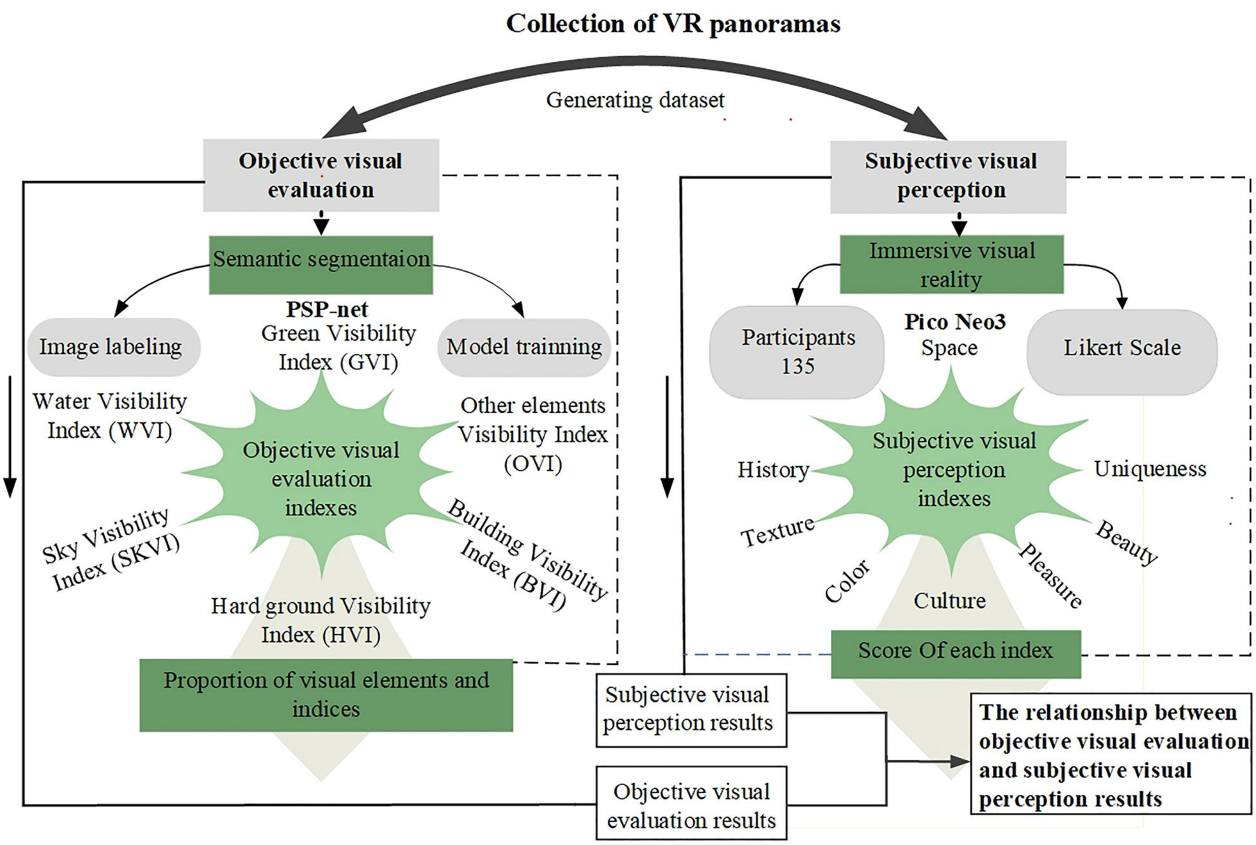

Fig. 2 Research framework

The evaluation questionnaire was composed of three parts. The first part was a personal information questionnaire, including gender, age, educational background, major, and number of visits to the railway. The second part was the visual perception questionnaire, encompassing a total of 40 Table 2 items. A 7-point Likert scale was used to rate each item from 1 (completely disagree) to 7 (completely agree). The visual perception score of each heritage site was the mean value of the overall index scores. The third part was a post-evaluation questionnaire, which consisted of the presence and immersion, and dizziness (IVR disease). There were two items in the questionnaire: (1) "What do you think of the presence and immersion of the IVR viewing experiment?" and (2) "Do you feel dizzy?".

Procedure

Participants were recruited through posters and social media platforms (Line). The inclusion criteria were no significant mood swings and cognitive or psychiatric disorders. The 135 participants (53 males and 82 females), aged between 15 and 57 years, were students and faculty from the College of Landscape and Horticulture, the College of Materials Science and Engineering, the College of Biodiversity, and other colleges of Southwest Forestry University. A small gift was presented to each participant after the experiment was completed as a token of appreciation.

The IVR viewing experiment was conducted at the 551 Workshop, Building A, Southwest Forestry University, from November 21 to 30, 2022. The experimental equipment consisted of a display screen, an Asus Notebook PC G512L, a PICO Neo3VR Head-mounted Display, and a free-rotating chair. The Pico Neo 3 is a Head-mounted Display that can read IVR panoramas directly. First, participants went to the information desk at the entrance of the workshop to complete the personal information questionnaire. In the meantime, a researcher explained the experiment's purpose and procedure. In addition, participants were informed that they could withdraw from the experiment at any time if they felt uncomfortable. To prevent the negative side effects (e.g., IVR sickness) from long exposure to IVR scenes and fatigue's influence on the score, the panoramas of 120 heritage sites were divided into 3 groups by an arithmetic progression, with each group experiencing 40 heritage sites. To ensure

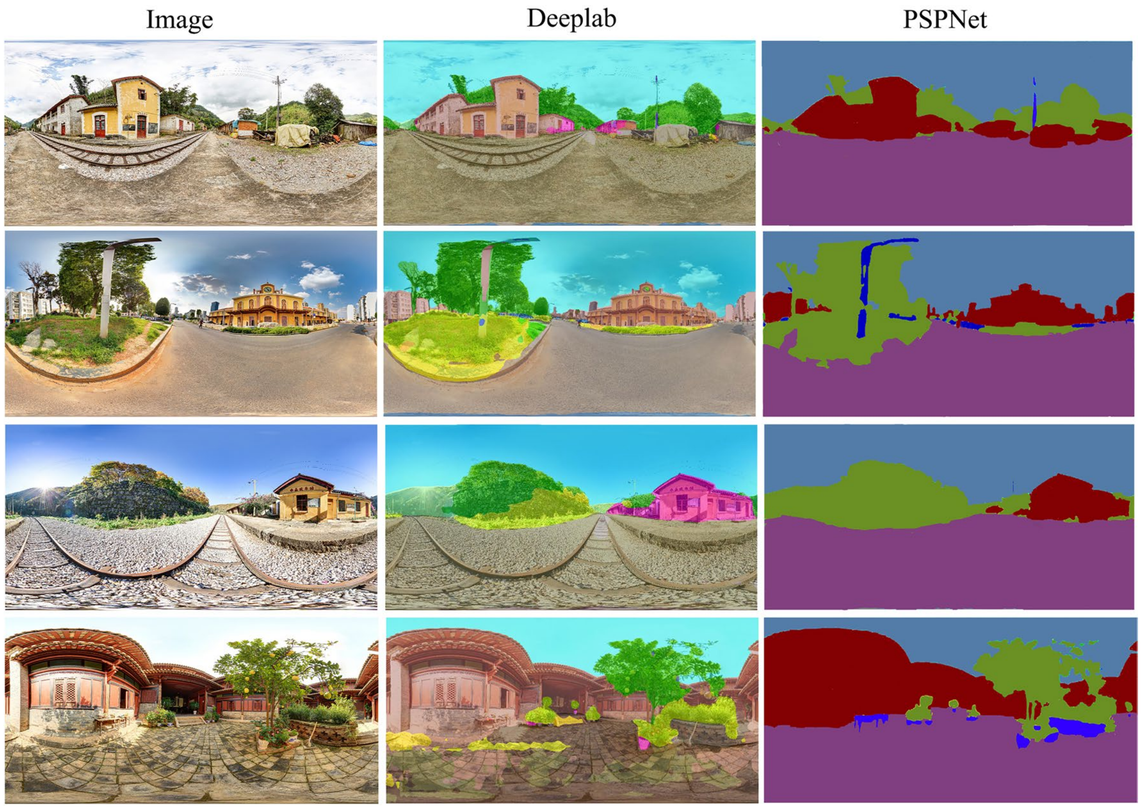

Fig. 3 Comparison of the segmentation effects between DeepLab and PSPNet on the railway heritage sites

Table 1 Description of the objective indices

Dimension	Indices	Indices description
Natural	GVI	The ratios of vegetation pixels (tree, grass, plant, mountain/hill) to the total pixels in the image
	WVI	The ratios of water pixels to the total pixels in the image
	SKVI	The ratios of sky pixels to the total pixels in the image
Artificial	HVI	The ratios of hard ground pixels (floor, road, sidewalk, earth/ground, rock) to the total pixels in the image
	BVI	The ratios of buildings pixels to the total pixels in the image
	OVI	The ratios of other elements pixels (person, car/van/minibike/boat /bus/truck/train, fence/wall, column/signboard/awning/streetlight /airplane/pole) to the total pixels in the image

Table 2 Rating scale for visual perception

Parameter	Description	Scale						
Space	Here is an open space with a wide view	1	2	3	4	5	6	7
Color	This heritage site is rich in color	1	2	3	4	5	6	7
Texture	The texture of this heritage site is clear	1	2	3	4	5	6	7
Uniqueness	This is a unique heritage site	1	2	3	4	5	6	7
Culture	This heritage site is rich in a cultural atmosphere	1	2	3	4	5	6	7
History	This heritage site has a long history	1	2	3	4	5	6	7
Beauty	This heritage site is beautiful	1	2	3	4	5	6	7
Pleasure	This heritage site gives me pleasure	1	2	3	4	5	6	7

comfortable viewing of the panorama, the head-mounted display was adjusted for participants. The participants could look around and score according to the description by the researcher. After viewing all the heritage sites, the participants needed to complete the post-evaluation questionnaire at the information desk and receive a gift. In order to eliminate potential interference, IVR viewing was performed with silent conditions. The whole experiment took approximately 30–35 min (Fig. 4).

Regression analysis between the objective and subjective visual evaluation

The stepwise multiple linear regression models [Eq. (1)] [43] were carried out in SPSS 25 to investigate the relationship between the objective and subjective visual evaluation. We first assumed that there was a linear relationship between the 9-category visual perception (space, color, texture, uniqueness, culture, history, beauty, pleasure) scores and 6 physical components (GVI, WVI, SKVI, HVI, BVI, OVI). Then, a certain linear regression model was used to fit the data of the variable, and the regression equation was obtained by determining the parameters. Since stepwise regression allows for the construction of regression models from a set of candidate variables, the system can automatically identify the influential variables, which helps to eliminate independent variables without significance and calculate an "optimal" regression equation for data with many variables that may not be entirely independent.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_p x_p + \varepsilon \tag{1}$$

where Y is the dependent variable; $x_1, x_2 \ldots x_p$ are the independent variables;

$\beta_0, \beta_1, \beta_2 \ldots \beta_p$ are the parameters; ε is random component (the rest of the model).

Results

Objective visual evaluation results

There are 16 successfully segmented elements (trees, grass, plant, mountain/hill, water, sky, floor, road, sidewalk, earth/ground, rock, buildings, person, car/van/minibike/ boat/bus/truck/train, fence/wall, column/signboard/ awning/streetlight/ pole) of the heritage sites. The visual element ratio of the sky accounts for the highest proportion of 33%, followed by earth/ground and trees,

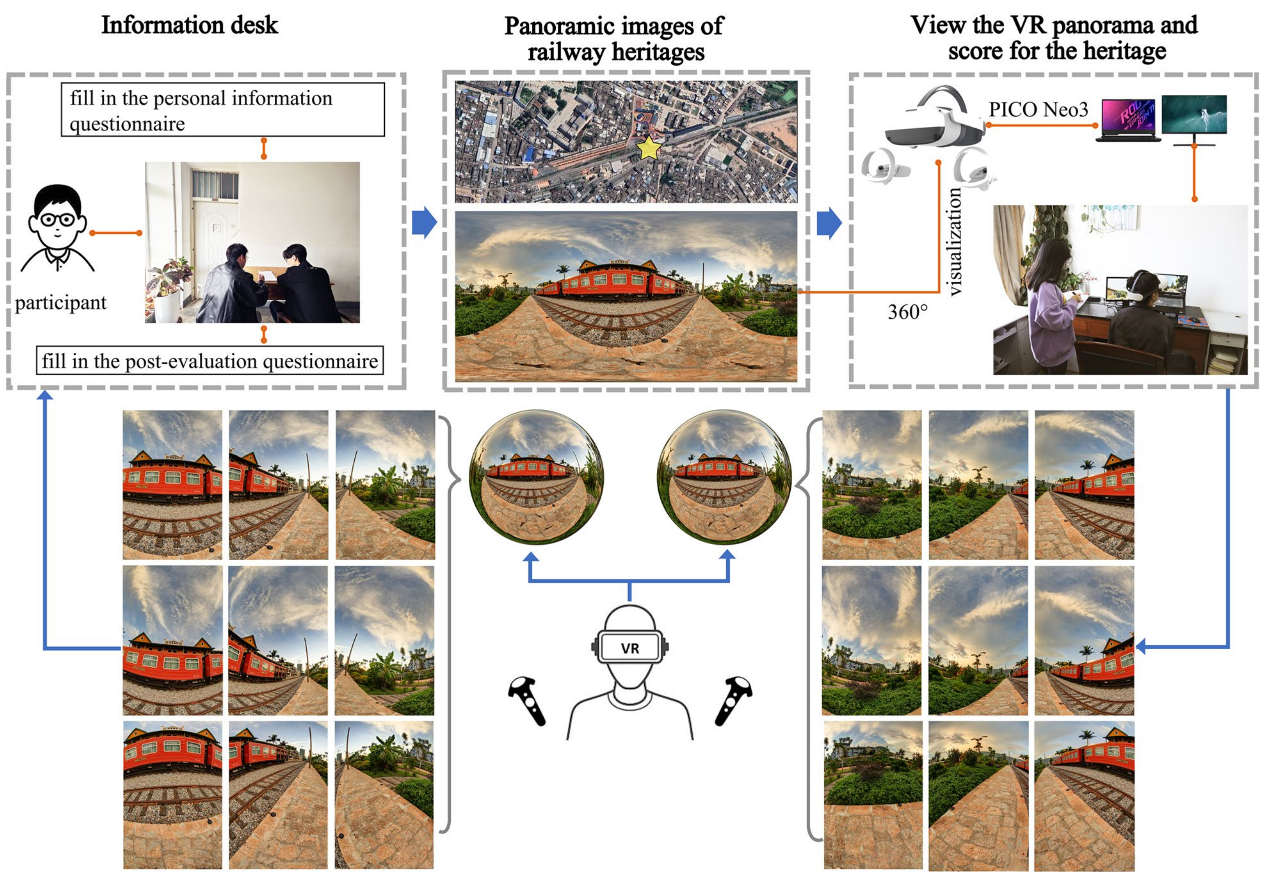

Fig. 4 The procedure of subjective visual perception

accounting for 22% and 10%, respectively. The visual element ratios of the floor, rock, buildings, grass, plant, mountain/hill, road, sidewalk, and fence/wall range from 1 to 10%, while the four elements of water, person, car and column account for less than 1%. To explore the objective visual characteristics of the heritage sites, 16 visual elements are classified into six indicators: GVI, WVI, SKVI, HVI, BVI and OVI (Fig. 5). We discover that the total proportion of SKVI (33%) and HVI (43%) is 78%, which forms the leading skeleton of the heritage landscape. GVI and BVI account for 16% and 7%, respectively, as secondary visual elements. OVI accounts for 1%, and WVI accounts for less than 1%.

We compared the objective characteristics of 120 heritage sites based on the pixel ratios of six objective visual indices in panoramic images. In most heritage sites, the geographic distributions of SKVI and HVI change mildly, while the BVI and GVI are noticeably unequal. In addition, the geographic distribution of WVI and OVI only exists in some special heritage sites. GVI consists of trees, grass, plant and mountain/hill, which is the main index of heritage sites 17, 26, 33, 79, 100, 111 and 113. WVI is predominantly found in bridges and partly in public service and productive heritage sites. There is no distribution of WVI in stations and residential heritage sites. OVI is mainly distributed in residential, public service and productive heritage sites (Fig. 6).

Subjective visual perception results

A total of 5400 responses (Likert scales) were collected for the railway heritage sites (45 participants × 40 sites × 3 groups = 5400 responses). Figure 7a shows the sum scores (SUMS) of the average scores of eight visual perception indices among 120 heritage sites. On the whole, SUMS of the heritage sites are relatively high, with the highest score being 48.77 and appearing at site 95, and the lowest score being 27.68 at site 19. According to the visual perception score bands, the 120 industrial heritage sites can be classified as 39 high-score sites (SUMS > 40), 66 medium-score sites (32 < SUMS ≤ 40), and 15 low-score sites (SUMS ≤ 32) (Fig. 8). The average scores of 87.5% of heritage sites are greater than 32 and the average scores of each perception index range exceeded 3, indicating

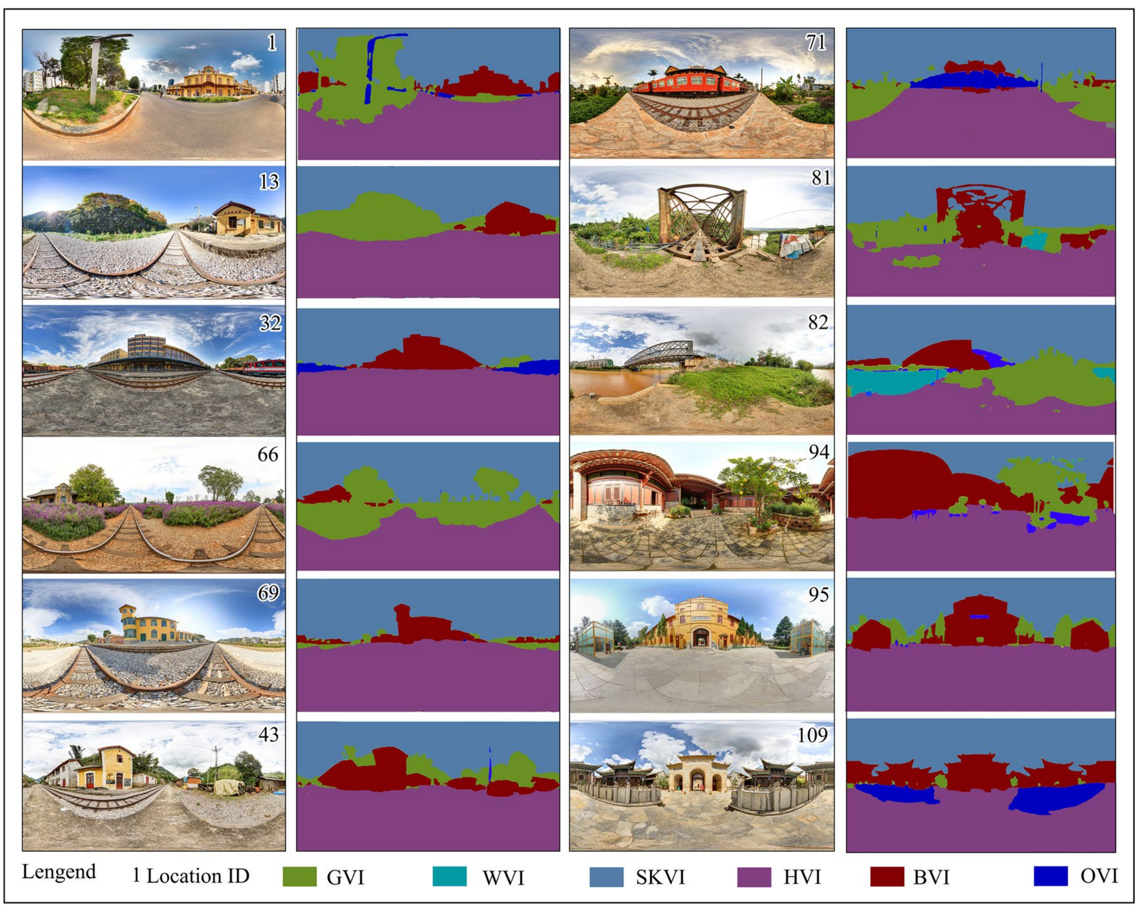

Fig. 5 Panoramic images and their semantic segmentation results

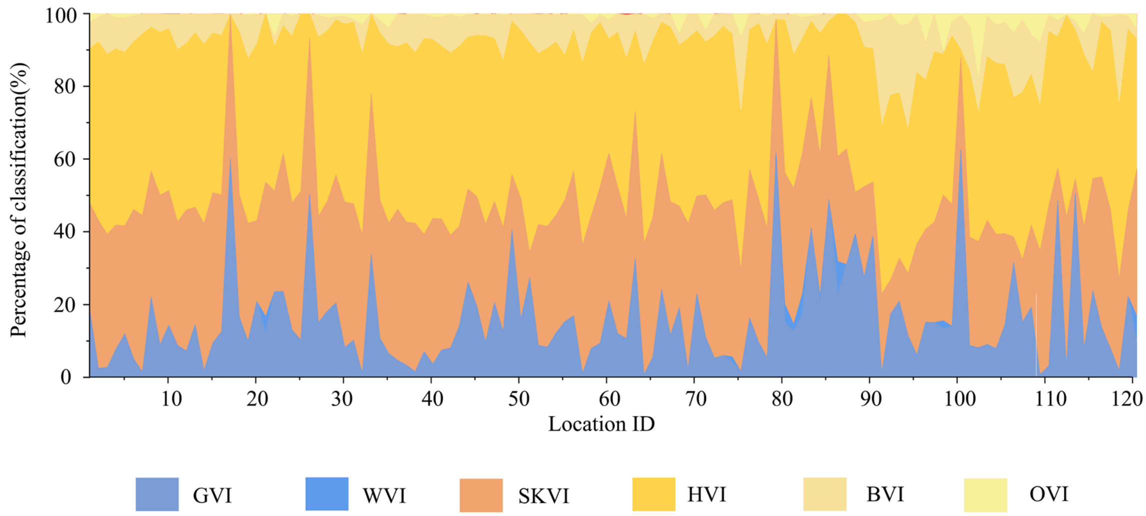

Fig. 6 Proportion and evaluation of visual indices

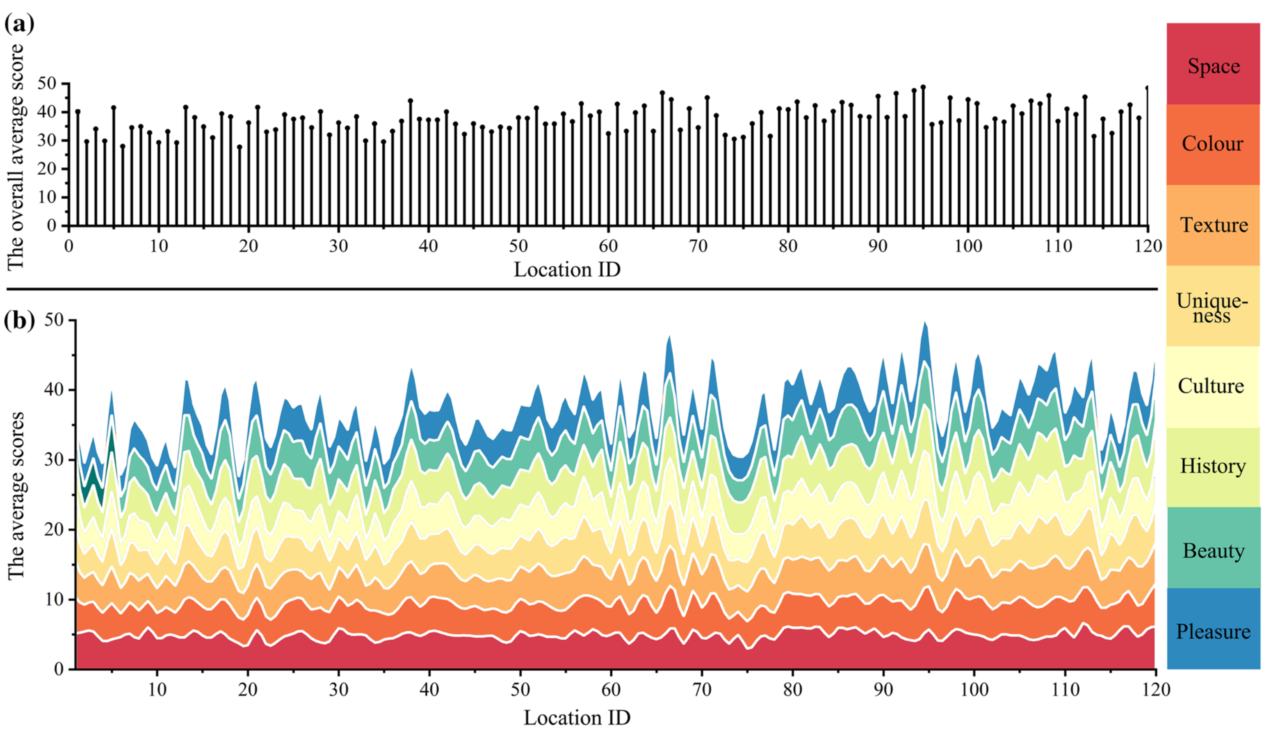

Fig. 7 **a** Overall average scores, **b** visual perception results visualized as a stream graph

that the majority of participants rated these heritage sites positively [44].

Figure 7b shows the results of the quantitative analysis of the eight visual perception indices in 120 heritage sites. In general, the average score of history is the highest at 5.00, with a maximum value of 6.51 at site 28 and a minimum value of 3.46 at site 33. We observed that the average score of pleasure is the lowest at 4.52, with a maximum value of 6.20 at site 120 and a minimum value of 3.15 at site 6. The average scores of space, color, texture, uniqueness, culture, and beauty are 4.93, 4.58, 4.75, 4.76, 4.60, and 4.58, respectively. The average scores of all indices ranged from 4.58 to 5.00, indicating that these

Fig. 8 Image samples of the railway heritage sites were predicted to have **a** high scores, **b** medium scores, and **c** low scores for visual perception

eight indices are all important to the visual quality of the heritage landscape.

The analysis of the post-evaluation questionnaire after viewing all the heritage sites shows that 95.6% of the participants thought that the presence and immersion of the IVR viewing experiment were strong, indicating that the immersive IVR experience provides people with a high-quality sense of presence. In addition, 10.4% of the participants experienced strong dizziness, 31.9% reported moderate dizziness, and 57.7% reported no symptoms. The findings regarding the occurrence of dizziness in virtual reality (VR) experiences demonstrate consistency across various environments and conditions of immersive virtual reality (IVR) usage, encompassing activities such as nature walks, stationary immersion, and active engagement within immersive environments [45, 46]. Therefore, IVR disease is still a problem that needs to be solved in the future when using IVR to conduct immersive evaluation research. In this experiment, although some participants felt dizzy, they were able to complete the experiment well after a brief rest.

Regression analysis results

The stepwise multiple regression analysis was used to determine the relationship between the ratios of objective visual evaluation and scores of subjective visual perceptions. We constructed nine stepwise regression models, corresponding to the nine categories (space, color, texture, uniqueness, culture, history, beauty, pleasure, and SUMS) with six objective indices. The Durban–Watson values of the nine models are 1.596, 1.984, 1.874, 2.070, 1.955, 2.021, 1.974, 2.066, and 2.011, respectively, which means that the data meet independence requirements. In all these models, the variance inflation factor (VIF) is less than 2, indicating that there is no potential multicollinearity problem. The residual histograms indicate that the residuals basically conform to a normal distribution. The results show that regression models can be built using these variables [34].

In the nine stepwise regression models, GVI, WVI, SKVI, BVI, and OVI are the explanatory variables, while HVI is the excluded variable. As shown in Fig. 9, GVI is positively correlated with beauty and pleasure, which is similar to the previous GVI analysis results of streets and parks [47]. WVI has a significant correlation with the scores of subjective visual perceptions, implying that water played a significant role in the positive perception of the heritage sites. The score of space increases with SKVI and WVI. The BVI is positively correlated with texture, uniqueness, culture and history, indicating that the railway buildings can be considered as an important representation of heritage characteristics [48]. In general, although the SKVI and HVI account for a higher proportion, they have little effect on the sum scores of the average scores of eight visual perception indices. GVI, WVI and BVI have a positive correlation with the sum scores, implying that vegetation, water and buildings play a significant role in the perception of visual quality.

Discussion

The significance of railway-related landscape planning and design has grown in importance within the broader context of environmental development due to the rise in the reuse and tourism of heritage railways. However, a planning approach that neglects public visual perception and solely focuses on top-down decision-making could potentially lead to the destruction of railway heritage sites instead of facilitating their efficient reuse [49, 50]. This study developed a visual evaluation and perception framework integrating computer vision (CV) and virtual reality (IVR) for quantifying the contributing visual elements and measuring the visual perception of the heritage sites. The results could provide important support for

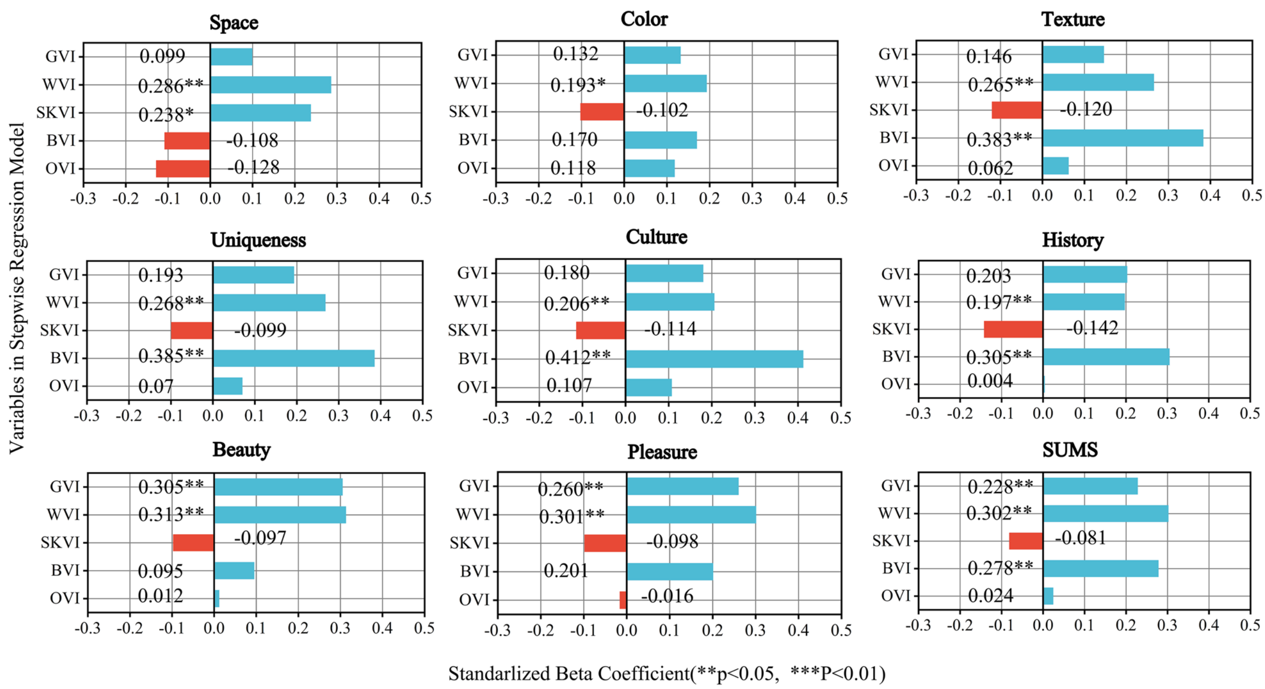

Fig. 9 Results of the stepwise multiple regression analysis between the objective visual evaluation and subjective perception scores

the management and development of the heritage sites of the Yunnan-Vietnam Railway (Yunnan section).

The creation of a semantic segmentation dataset for railway heritage sites enables a quantitative analysis of the landscape components within extensive railway heritage sites, thereby facilitating the identification and understanding of visual attributes associated with railway landscapes. We discovered that the total proportion of SKVI (33%) and HVI (43%) is 78%, which forms the leading skeleton of the heritage landscape. This is related to the panorama feature [51]. GVI and SKVI accounted for 16% and 7%, respectively, as secondary visual features. OVI accounted for 1%, and water (WVI) accounted for less than 1%. The results suggest that the semantic segmentation model (e.g., PSPNet) can be used to batch process large-scale linear heritage landscape photos, which has certain applicability [33]. In addition, the research focuses on linear heritage, which is conducive to other large-scale linear heritage research.

Previous studies have confirmed that IVR panoramas are more conducive to landscape visual characterization than traditional 2D images in terms of immersion and presence [45]. Head-mounted displays were used to realize the perception of the heritage landscape from the normal human perspective, bringing a high-quality sense of presence. This method allows for a larger field of view and has become an important aid for visual evaluation. The average scores of 87.5% of heritage sites are greater than 32, and the average scores of each perception index range exceeded 3, indicating that the majority of participants rated these heritage sites positively [44]. According to the visual perception score bands, the 120 industrial heritage sites can be classified as 39 high-score sites (SUMS > 40), 66 medium-score sites (32 < SUMS ≤ 40), and 15 low-score sites (SUMS ≤ 32). In the process of heritage management and development, it is possible to implement cluster management and to take measures for the different score clusters. In the post-evaluation phase, although the participants greatly affirmed the presence and immersion of the IVR panorama, some of them suffered from dizziness, which can also be called "IVR" disease. This is also a problem with most relevant experiments [52], indicating that further research is needed in this area.

According to the results of the stepwise multiple regression analysis, we discover that GVI is positively correlated with beauty and pleasure, which is similar to the results of GVI analyses conducted at the street and park levels [47]. Thus, the impact of GVI on visual perception is applicable not only to street and park landscapes but also to railway industry heritage landscapes. WVI has a significant correlation with the scores of subjective visual perceptions, which differs from the findings of Luo et al. [33]. This is related to water quality. The Yunnan-Vietnam Railway is close to the Nanpan River and Nanxi River, and there are some highland lakes along the route. In the process of fieldwork, we found that the water has a high quality, which form beautiful scenery. Therefore, good water quality can improve people's visual experiences, which is also supported by previous research [53]. The score of the space increases with SKVI and WVI, indicating that the higher the proportion of sky and water, the more open the space is. The BVI is positively correlated with texture, uniqueness, culture, and history. This is in line with the selection of indicators related to world heritage [54]. Although the SKVI and HVI accounted for a higher proportion, they had little effect on the sum scores, while the GVI, WVI, and BVI had a positive correlation with the sum scores. Compared to buildings, vegetation, and water, the sky and ground mostly lack variation and mostly exist as a background. The results of regression analyses contribute to the advancement of the humanization and refinement of the identification, evaluation, and management of objective landscape elements associated with subjective visual perception scores.

Although we demonstrate that IVR panoramic technology and semantic segmentation techniques can be used to analyze the visual features of railway heritage landscapes, there are still some problems to be solved. First, the semantic segmentation of images can be improved with further efforts. There are also inaccuracies in the recognition of elements in different states, such as historical walls being recognized as grass because of the growth of grass on them and dead branches in winter not being recognized as trees. Second, although we used similar parameters and conditions for filming to maintain data consistency, there are still inconsistency problems caused by the large number and span of the railway heritage sites. Finally, panorama-based evaluation methods can be more diverse, such as using an immersive sensing device or nonimmersive sensing equipment, or a combination of both. Immersion in IVR provides a strong sense of presence, while nonimmersion in IVR allows more people to take part in the experiment remotely. The limitations and challenges in the study will also be topics to address in future research. Through a standardized panoramic data processing process, the research results will not only promote replication in different locations but also allow iterative updates within the same location, and it will help analyze the dynamics of landscape change over time, thus improving the digitization and refinement of heritage spatial management, which is useful for heritage planners, environmental managers, and railway researchers.

Conclusion

This study developed an "objective + subjective" visual evaluation and perception framework integrating CV and IVR to assess the visual quality of industrial heritage sites along the Yunnan-Vietnam Railway (Yunnan Section), which is a novelty in this domain. We generated a semantic segmentation dataset based on 120 IVR panoramic images of the railway heritage landscape. Using this dataset and the CV algorithm, the railway heritage and the surrounding landscape can be automatically and effectively analyzed, overcoming the shortcomings of existing techniques. According to the visual perception score bands, the 120 industrial heritage sites can be classified as 39 high-score sites, 66 medium-score sites, and 15 low-score sites. The participants' overall level of preference regarding heritage landscapes can be a significant factor in deciding on the reuse ways, which facilitates the formation of a bottom-up planning scheme that incorporates public visual perception. We also identified the visual elements that strongly influence human visual perception by using the multivariate stepwise linear regression models. In general, although the sky and hard ground accounted for a higher proportion, they had little effect on the sum scores, while the vegetation, water, and buildings played a significant role in the perception of visual quality. Furthermore, we investigated the possible reasons for the correlation results and compared them with those of related studies. Thus, railway heritage planners and managers can consider adding dense vegetation, water, and buildings to these settings to build high-quality heritage environments. The results and proposed framework can help researchers, planners, and government departments clarify the visual quality to scientifically specify bottom-up planning and management solutions for railway industrial heritage sites, which are beneficial to railway-related landscapes and environments.

Appendix A

See Table 3.

Table 3 IDs and names of railway industrial heritage sites

ID	Name of the heritage site	ID	Name of the heritage site	ID	Name of the heritage site
1	Kunmingbei	41	Luoshuidong	81	The French railrway bridge at Xiaolongtan
2	Heituao	42	Gegu	82	the Muhuaguo bridge
3	Niujiezhuang	43	Luogu	83	the Yulinshan seven-hole bridge
4	Xiaoxicun	44	Tingtang	84	the Renzhecun three-hole bridge
5	Xizhuang	45	Bodujing	85	the eight-hole bridge
6	Chenggong	46	Chongzhuang	86	theWujiazhai bridge
7	Wangjiaying	47	Wantang	87	the Baizhai bridge
8	Sanjiacun	48	Baizhai	88	the Laofanzhai bridge
9	Shuitang	49	Baiheqiao	89	the China-Vietnam railway bridge
10	Yangzonghai	50	Lahadi	90	the Lu Han Residence
11	Fengmingcun	51	Dashutang	91	the Hu Zhiming former residence in Kunming
12	Kebaocun	52	Laofanzhai	92	the Zhou family courtyard
13	Shuijingpo	53	Nanxi	93	the Wang family courtyard
14	Yiliang	54	Majie	94	the Yuan Jiagu former residence in Shiping
15	Yangjiezi	55	Mahuangbao	95	the former site of the army officials school in yunnan
16	Goujiezi	56	Shanyao	96	the former site of Ganmei hospital
17	Dishui	57	Hekou	97	the former site of French hospital
18	Xujiadu	58	Yuguopu	98	the 1909 square
19	Lufengcun	59	Jiangshuidi	99	the French garden
20	Nuozu	60	Jijie	100	the site of the French consular office
21	Dashatian	61	Datianshan	101	the former site of Customs and Excise department
22	Xier	62	Wulichong	102	the French prison
23	Xiaohekou	63	Nanyingzhai	103	the Colombo Lu Si western travel
24	Panxi	64	Linan	104	the former site of estuary Customs office of flood control
25	Reshuitang	65	Jianshui	105	the Customs site and ancient fort in Hekou
26	Xicheyi	66	Xianghuiqiao	106	the former site of post office in Hekou
27	Lalihei	67	Tuanshan	107	the former site of the Gebishi railway company
28	Xunjiansi	68	Xiapochu	108	the Qihe Tower
29	Denglongshan	69	Baxin	109	the Chen family ancestral hall
30	Xiaolongtan	70	Renshoucun	110	the former site of central electrical equipment factory
31	Shilicun	71	Shiping	111	the water works in Kunming
32	Kaiyuan	72	Songcun	112	the coal mine in Xiaolongtan
33	Yulinshan	73	Baoxiu	113	the former site of the power plant in Kaiyuan
34	Zhumashao	74	Huogudu	114	the Nanqiao power plant
35	Data	75	Gejiu	115	the baofenglong firm
36	Dazhuang	76	Renhecun	116	the Yunxi Datun concentrator
37	Caoba	77	Baishachong	117	the Marag mine in Gejiu
38	Bisezhai	78	Mengzi	118	The concentrator and bell tower in Gejiu
39	Heilongtan	79	the Goujie bridge	119	the former site of Yunnan Tin company
40	Zhicun	80	The Lufengcun bridge	120	the Zhujia Garden

Abbreviations

CV	Computer vision
IVR	Immersive virtual reality
2D	Two-dimensional
ID	Identity document
GVI	Green visibility index
WVI	Water visibility index
SKVI	Sky visibility index
HVI	Buildings visibility index
OVI	Other elements visibility index
SUMS	The sum scores of the average scores of eight visual perception indices
VIF	The variance inflation factor

Acknowledgements
The authors express our acknowledgement to all participants for their valuable assessment.
We are grateful to our team for data collection and to the people who gave us directions and explanations in the field.

Author contributions
All the experiments were designed by WY and WJ. The data were acquired and analyzed by WY, WS, PY, CC and LC. The manuscript was written by WY, and revised by WJ. All authors read and approved the final manuscript.

Funding
This research was funded by the National Natural Science Foundation of China (Project name: Study on Silicon biomineralization in bamboos; project code: 31460169) as well as the Scientific Research Foundation of Yunnan Educational Committee (Project name: Study on the landscape composition of the Yunnan-Vietnam Railway in the context of National Cultural Park; project code: 2022Y613).

Availability of data and materials
The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 12 July 2023 Accepted: 12 January 2024
Published online: 02 February 2024

References

1. Rovelli R, Senes G, Fumagalli N, Sacco J, de Montis A. From railways to greenways: a complex index for supporting policymaking and planning. A case study in Piedmont (Italy). Land Use Policy. 2020;99:104835. https://doi.org/10.1016/j.landusepol.2020.104835.
2. Blancheton B, Marchi JJ. The three systems of rail tourism: French case. Tour Manag Perspect. 2013;5:31–40. https://doi.org/10.1016/j.tmp.2012.09.008.
3. Sang K, Fontana GL, Piovan SE. Assessing railway landscape by ahp process with gis: a study of the Yunnan-Vietnam railway. Remote Sens. 2022;14:603. https://doi.org/10.3390/rs14030603.
4. Eizaguirre-Iribar A, Grijalba O. A methodological proposal for the analysis of disused railway lines as territorial structuring elements: the case study of the Vasco-Navarro railway. Land Use Policy. 2020;91: 104406. https://doi.org/10.1016/j.landusepol.2019.104406.
5. Krause CL. Our visual landscape: Managing the landscape under special consideration of visual aspects. Landsc Urban Plan. 2001;54:239–54. https://doi.org/10.1016/S0169-2046(01)00139-6.
6. Jeon JY, Jo HI. Effects of audio-visual interactions on soundscape and landscape perception and their influence on satisfaction with the urban environment. Build Environ. 2020;169: 106544. https://doi.org/10.1016/j.buildenv.2019.106544.
7. Qi Y, Chodron Drolma S, Zhang X, Liang J, Jiang H, Xu J, Ni T. An investigation of the visual features of urban street vitality using a convolutional neural network. Geo-Spat Inf Sci. 2020;23:341–51. https://doi.org/10.1080/10095020.2020.1847002.
8. Ma B, Hauer RJ, Xu C, Li W. Visualizing evaluation model of human perceptions and characteristic indicators of landscape visual quality in urban green spaces by using nomograms. Urban For Urban Green. 2021;65: 127314. https://doi.org/10.1016/j.ufug.2021.127314.
9. Blečić I, Cecchini A, Trunfio G A. Towards automatic assessment of perceived walkability. In Computational Science and Its Applications–ICCSA 2018: 18th International Conference, Melbourne, VIC, Australia, July 2–5, 2018, Proceedings, Part III 18 (pp. 351–365). Springer International Publishing.
10. Li X, Wang X, Jiang X, Han J, Wang Z, Wu D, et al. Prediction of riverside greenway landscape aesthetic quality of urban canalized rivers using environmental modeling. J Clean Prod. 2022;367: 133066. https://doi.org/10.1016/j.jclepro.2022.133066.
11. Li X, Li L, Wang X, Lin Q, Wu D, Dong Y, Han S. Visual quality evaluation model of an urban river landscape based on random forest. Ecol Indic. 2021;133: 108381. https://doi.org/10.1016/j.ecolind.2021.108381.
12. Hong S, Kwak S, Han B. Weakly supervised learning with deep convolutional neural networks for semantic segmentation: Understanding semantic layout of images with minimum human supervision. IEEE Signal Process Mag. 2017;34:39–49. https://doi.org/10.1109/MSP.2017.2742558.
13. Li L, Li X, Jiang L, Su X, Chen F. A review on deep learning techniques for cloud detection methodologies and challenges. Signal Image Video Process. 2021;15:1527–35. https://doi.org/10.1007/s11760-021-01885-7.
14. Ki D, Lee S. Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning. Landsc Urban Plan. 2021;205: 103920. https://doi.org/10.1016/j.landurbplan.2020.103920.
15. Yang Z, Yu H, Feng M, Sun W, Lin X, Sun M, et al. Small object augmentation of urban scenes for real-time semantic segmentation. IEEE Trans Image Process. 2020;29:5175–90. https://doi.org/10.1109/TIP.2020.2976856.
16. Liu B, Fan R. Quantitative analysis of the visual attraction elements of landscape space. J Nanjing For Univ. 2014;38:149–52 (**(in Chinese)**).
17. Li X, Zhang C, Li W, Ricard R, Meng Q, Zhang W. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For Urban Green. 2015;14:675–85. https://doi.org/10.1016/j.ufug.2015.06.006.
18. Yin L, Wang Z. Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. Appl Geogr. 2016;76:147–53. https://doi.org/10.1016/j.apgeog.2016.09.024.
19. Tang J, Long Y. Measuring visual quality of street space and its temporal variation: methodology and its application in the Hutong area in Beijing. Landsc Urban Plan. 2019;191: 103436. https://doi.org/10.1016/j.landurbplan.2018.09.015.
20. Jahani A, Saffariha M. Aesthetic preference and mental restoration prediction in urban parks: an application of environmental modeling approach. Urban For Urban Green. 2020;54: 126775. https://doi.org/10.1016/j.ufug.2020.126775.
21. Pathak R, Saini A, Wadhwa A, Sharma H, Sangwan D. An object detection approach for detecting damages in heritage sites using 3-D point clouds and 2-D visual data. J Cult Herit. 2021;48:74–82. https://doi.org/10.1016/j.culher.2021.01.002.
22. Zube EH, Sell JL, Taylor JG. Landscape perception: research, application and theory. Landscape Plann. 1982;9:1–33. https://doi.org/10.1016/0304-3924(82)90009-0.
23. Qin X, Fang M, Yang D, Wangari VW. Quantitative evaluation of attraction intensity of highway landscape visual elements based on dynamic perception. Environ Impact Assess Rev. 2023;100: 107081. https://doi.org/10.1016/j.eiar.2023.107081.
24. Junge X, Schüpbach B, Walter T, Schmid B, Lindemann-Matthies P. Aesthetic quality of agricultural landscape elements in different seasonal stages in Switzerland. Landsc Urban Plan. 2015;133:67–77. https://doi.org/10.1016/j.landurbplan.2014.09.010.
25. Sun L, Wang J, Yang K, Wu K, Zhou X, Wang K, Bai J. Aerial-PASS: panoramic annular scene segmentation in drone videos. In 2021 European Conference on Mobile Robots (ECMR) (pp. 1–6). IEEE. doi: https://doi.org/10.1109/ECMR50962.2021.9568802.

26. Sun L, Shao H, Li S, Huang X, Yang W. Integrated application of eye movement analysis and beauty estimation in the visual landscape quality estimation of urban waterfront park. Intern J Pattern Recognit Artif Intell. 2018;32:1856010. https://doi.org/10.1142/S0218001418560104.
27. De Paolis LT, Chiarello S, Gatto C, Liaci S, De Luca V. Virtual reality for the enhancement of cultural tangible and intangible heritage: The case study of the Castle of Corsano. Digit Appl Archaeol Cult Herit. 2022;27: e00238. https://doi.org/10.1016/j.daach.2022.e00238.
28. Brivio E, Serino S, Negro Cousa E, Zini A, Riva G, De Leo G. Virtual reality and 360 panorama technology: a media comparison to study changes in sense of presence, anxiety, and positive emotions. Virtual Real. 2021;25:303–11. https://doi.org/10.1007/s10055-020-00453-7.
29. Verhulst I, Woods A, Whittaker L, Bennett J, Dalton P. Do VR and AR versions of an immersive cultural experience engender different user experiences? Comput Human Behav. 2021;125: 106951. https://doi.org/10.1016/j.chb.2021.106951.
30. Arnold D, Day A, Glauert J, Haegler S, Jennings V, Kevelham B, Laycock R, et al. Tools for populating cultural heritage environments with interactive virtual humans. In Open Digital Cultural Heritage Systems Conference. 2008;25:25.
31. Dong L, Jiang H, Li W, Qiu B, Wang H, Qiu W. Assessing impacts of objective features and subjective perceptions of street environment on running amount: a case study of Boston. Landsc Urban Plan. 2023;235: 104756. https://doi.org/10.1016/j.landurbplan.2023.104756.
32. Gao Y, Zhang T, Zhang W, Meng H, Zhang Z. Research on visual behavior characteristics and cognitive evaluation of different types of forest landscape spaces. Urban For Urban Green. 2020;54: 126788. https://doi.org/10.1016/j.ufug.2020.126788.
33. Luo J, Zhao T, Cao L, Biljecki F. Semantic Riverscapes: Perception and evaluation of linear landscapes from oblique imagery using computer vision. Landsc Urban Plan. 2022;228: 104569. https://doi.org/10.1016/j.landurbplan.2022.104569.
34. Li Y, Yabuki N, Fukuda T. Measuring visual walkability perception using panoramic street view images, virtual reality, and deep learning. Sustain Cities Soc. 2022;86: 104140. https://doi.org/10.1016/j.ufug.2021.127314.
35. Kim SN, Lee H. Capturing reality: Validation of omnidirectional video-based immersive virtual reality as a streetscape quality auditing method. Landsc Urban Plan. 2022;218: 104290. https://doi.org/10.1016/j.landurbplan.2021.104290.
36. Jo HI, Jeon JY. Perception of urban soundscape and landscape using different visual environment reproduction methods in virtual reality. Appl Acoust. 2022;186: 108498. https://doi.org/10.1016/j.apacoust.2021.108498.
37. Hu CB, Zhang F, Gong FY, Ratti C, Li X. Classification and mapping of urban canyon geometry using Google Street View images and deep multitask learning. Build Environ. 2020;167: 106424. https://doi.org/10.1016/j.buildenv.2019.106424.
38. Li W, Zhai J, Zhu M. Characteristics and perception evaluation of the soundscapes of public spaces on both sides of the elevated road: a case study in Suzhou. China Sustain Cities Soc. 2022;84: 103996. https://doi.org/10.1016/j.scs.2022.103996.
39. Kang N, Liu C. Towards landscape visual quality evaluation: methodologies, technologies, and recommendations. Ecol Indic. 2022;142: 109174. https://doi.org/10.1016/j.ecolind.2022.109174.
40. Wang Y, Du J, Kuang J, Chen C, Li M, Wang J. Two-scaled identification of landscape character types and areas: a case study of the Yunnan-Vietnam Railway (Yunnan Section). China Sustain. 2023;15:6173. https://doi.org/10.3390/su15076173.
41. Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2881–2890).
42. Li C, Shen S, Ding L. Evaluation of the winter landscape of the plant community of urban park green spaces based on the scenic beauty esitimation method in Yangzhou China. PLoS ONE. 2020;15: e0239849. https://doi.org/10.1371/journal.pone.0239849.
43. Kolasa-Wiecek A. Stepwise multiple regression method of greenhouse gas emission modeling in the energy sector in Poland. J Environ Sci. 2015;30:47–54. https://doi.org/10.1016/j.jes.2014.09.037.
44. Jeon JY, Jo HI, Lee K. Potential restorative effects of urban soundscapes: personality traits, temperament, and perceptions of VR urban environments. Landsc Urban Plan. 2021;214: 104188. https://doi.org/10.1016/j.landurbplan.2021.104188.
45. Luo S, Shi J, Lu T, Furuya K. Sit down and rest: use of virtual reality to evaluate preferences and mental restoration in urban park pavilions. Landsc Urban Plan. 2022;220: 104336. https://doi.org/10.1016/j.landurbplan.2021.104336.
46. Calogiuri G, Litleskare S, Fagerheim KA, Rydgren TL, Brambilla E, Thurston M. Experiencing nature through immersive virtual environments: environmental perceptions, physical engagement, and affective responses during a simulated nature walk. Front Psychol. 2018;8:2321. https://doi.org/10.3389/fpsyg.2017.02321.
47. Zhang F, Zhou B, Liu L, Liu Y, Fung HH, Lin H, Ratti C. Measuring human perceptions of a large-scale urban region using machine learning. Landsc Urban Plan. 2018;180:148–60. https://doi.org/10.1016/j.landurbplan.2018.08.020.
48. Cucco P, Maselli G, Nesticò A, Ribera F. An evaluation model for adaptive reuse of cultural heritage in accordance with 2030 SDGs and European Quality Principles. J Cult Herit. 2023;59:202–16. https://doi.org/10.1016/j.culher.2022.12.002.
49. Fagerholm N, Käyhkö N, Van Eetvelde V. Landscape characterization integrating expert and local spatial knowledge of land and forest resources. Environ Manage. 2013;52:660–82. https://doi.org/10.1007/s00267-013-0121-x.
50. Wartmann FM, Acheson E, Purves RS. Describing and comparing landscapes using tags, texts, and free lists: an interdisciplinary approach. Int J Geogr Inf Sci. 2018;32:1572–92. https://doi.org/10.1080/13658816.2018.1445257.
51. Orhan S, Bastanlar Y. Semantic segmentation of outdoor panoramic images. Signal Image Video Process. 2022;16:643–50. https://doi.org/10.1007/s11760-021-02003-3.
52. Somrak A, Humar I, Hossain MS, Alhamid MF, Hossain MA, Guna J. Estimating VR Sickness and user experience using different HMD technologies: An evaluation study. Future Gener Comput Syst. 2019;94:302–16. https://doi.org/10.1016/j.future.2018.11.041.
53. Li X, Chen WY, Hu FZY, Cho FHT. Homebuyers' heterogeneous preferences for urban green–blue spaces: a spatial multilevel autoregressive analysis. Landsc Urban Plan. 2021;216: 104250. https://doi.org/10.1016/j.landurbplan.2021.104250.
54. Kimura F, Ito Y, Matsui T, Shishido H, Kitahara I, Kawamura Y, Morishima A. Tourist participation in the preservation of world heritage–a study at Bayon temple in Cambodia. J Cult Herit. 2021;50:163–70. https://doi.org/10.1016/j.culher.2021.05.001.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.