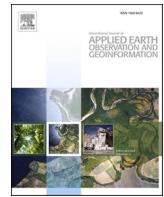


Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



Measuring cyclists' subjective perceptions of the street riding environment using K-means SMOTE-RF model and street view imagery

Qisheng Zeng, Zheng Gong, Songtai Wu, Caigang Zhuang, Shaoying Li *

School of Geography and Remote Sensing, Guangzhou University, Guangzhou 510006, China



ARTICLE INFO

Keywords:
Subjective perceptions
Street environment
Cyclists
Machine learning
K-means SMOTE
Street view imagery

ABSTRACT

Cyclists' willingness to ride is usually influenced by their subjective perception of the street riding environment. Measuring this perception is crucial for enhancing residents' willingness to ride. We propose an SSB framework (Public Security, Traffic Safety, Scenic Beauty) to quantify cyclists' subjective perception using street view imagery (SVI) and volunteer rating data. To address the issue of imbalanced class distribution in the volunteer rating data and enhance the model's ability to distinguish between positive and negative perception scenes, we employed a combination of the Kmeans Synthetic Minority Over-Sampling Technique (Kmeans-SMOTE) and the Random Forest (RF) classifier. The Kmeans SMOTE-RF model improved Area Under the Curve (AUC) by 0.327 for public safety, 0.2 for traffic safety, and 0.209 for scenic beauty compared to the RF model. Additionally, we incorporated Shapley Additive Explanations (SHAP) to examine how the visual features of SVI impact cyclists' subjective perception. Trees had a positive impact on all dimensions. Fence and sidewalk were key features for enhancing traffic safety perception, while roads positively affected public security and scenic beauty. These insights support urban planners in understanding the relationship between SVI features and cyclists' perceptions, aiding the design of cyclist-friendly street environments.

1. Introduction

Transportation carbon emissions constitute a substantial contributor to global carbon emissions, with the transportation sector accounting for approximately 24 % of the world's carbon emissions between 2000 and 2018 (Lin et al., 2021). Although the transportation sector accounts for only 10 % of China's total carbon emissions, China's rapid economic growth has led to a significant increase in both passenger and freight transport, making transportation one of the fastest growing sectors in China's carbon emissions (Zhang & Hanaoka, 2022). To mitigate this trend, relieving the dependence of cars on fossil fuels, optimizing the travel structure and increasing the proportion of public transportation and slow travel are crucial strategies for reducing transportation carbon emissions (Cervero & Sullivan, 2011).

Cycling is considered a sustainable mode of transportation, with cyclists emitting 84 % less carbon dioxide during travel compared to non-cyclists (Brand et al., 2021). Additionally, cycling can promote the use of public transportation, due to the fact that cycling play an important role in solving the problem of the first and last mile of public transportation (Guo & He, 2020). Because of these advantages of

cycling, many cities attempt to guide citizens to travel more by bicycle to reduce urban traffic carbon emissions (Ellison & Greaves, 2011; Li et al., 2021; Luan et al., 2020). Therefore, exploring the impact mechanism of cycling behavior is crucial to improving residents' willingness to ride.

The factors influencing cycling behavior are complicated. A substantial body of literature has looked into the association of cycling behavior and objective built environment, which is usually measured by the 5D framework (i.e., diversity, density, distance to transit, design and destination accessibility) (Ewing & Cervero, 2010). These studies mainly focused on the grid scale (Gao et al., 2021; Shen et al., 2018; Xu et al., 2019) and traffic analysis zone (TAZ) level (Chen & Ye, 2021; Ma et al., 2020; Tu et al., 2019). Some other studies have attempted to examine the suitability of physical urban environment for cycling at the street level. Various factors, such as street greening (Chen et al., 2020; Gao et al., 2021; Wang et al., 2020), street morphology (Meng & Zacharias, 2021), traffic conditions (Winters et al., 2013), as well as the quality and quantity of bicycle-related infrastructure (Aziz et al., 2018; Winters & Teschke, 2010), have been shown to play a substantial role in cycling suitability. These empirical findings have enriched the understanding of the impact of objective cycling environment on cycling behavior.

* Corresponding author.

E-mail address: lsy@gzhu.edu.cn (S. Li).

In addition to the objective cycling environment, the subjective perception of the street environment by cyclists is also an important factor affecting their willingness to ride. Measuring urban residents' subjective perceptions has long been an area of interest for urban researchers. Early studies primarily relied on some time-consuming, labor-intensive, and costly data collection methods such as interviews and questionnaires (Cresswell, 1992; Montello et al., 2003). Recently, the emergence of SVI and advances in machine learning have made it possible to capture the subjective perceptions of urban residents on a large scale. The widely used method for evaluating residents' subjective perceptions is the two-step approach. First, the subjective perceptions from a few volunteers are collected. Then, these data are combined with machine learning models to simulate a wide range of subjective perceptions. A recent study by Li et al. (2022) combined SVI with virtual reality panoramic images to measure visual walkability perceptions in volunteers. Zhang et al. (2018) developed a method for predicting residents' sense of place, which includes six dimension perceptions: safe, lively, wealthy, beautiful, depressing, and boring. Kruse et al. (2021) assessed human perceptions of urban playability using a ResNet18 deep learning model trained on labeled streetscape data. This study explored the relationship between speed limits, crime, the presence of trees, amusement parks, and playability. The combination of SVI and machine learning techniques has opened new avenues for understanding and modeling residents' subjective perceptions of the urban environment. However, few studies have been conducted to capture the subjective perceptions of cyclists and simulate their perception on the cycling environment at the street level.

To fill this gap, this study will propose a framework for evaluating the visual subjective perception of cyclists based on SVI data. We employ an interpretable machine learning model to investigate the influence of visual features within the streetscape on cyclists' subjective perceptions. To initiate the study, we develop a Streetview Rating Software and recruited 50 volunteers. These volunteers were requested to rate scenes portrayed in SVI from the viewpoint of cyclists, focusing on three dimensions: public security, traffic safety, and scenic beauty (SSB). In existing research on resident perception evaluation based on machine learning and SVI dataset, a limited volunteer datasets often led to imbalance in the classification categories within the data. Consequently, the trained models tended to overfit the majority class while neglecting the minority class. Few studies have effectively handled the challenges associated with small-scale volunteer data collection and accurate modeling. To solve this issue, we develop a machine learning approach that integrated the Kmeans-SMOTE oversampling method with a random forest classifier to model and predict cyclists' perceptions of the street environment. This approach not only mitigates the issue of overfitting caused by imbalanced categories in volunteer rating data but also enables personalized and accurate perception modeling for each individual. Finally, we employ the Shapley Additive Explanations (SHAP) and Spearman correlation index to investigate the relationship between the visual features of SVI and cyclists' subjective perceptions on the streets. The study explores how the visual features of SVI affect cyclists' subjective perception, providing valuable insights for urban planners to create more cyclist-friendly streets.

This paper is structured as follows: Section 2 reviews the relevant literature; Section 3 describes the study area, datasets, and the modeling methods; Section 4 presents the results and discusses the findings; Section 5 provides a conclusion for the research.

2. Related work

2.1. Human subjective perception measurement with SVI and machine learning method

SVI offers a valuable urban data source, covering half of the world's population (Goel et al., 2018). Machine learning advancements enable the extraction and quantification of visual features from extensive SVI

datasets, allowing urban researchers to analyze cities from a human-eye perspective. These visual features derived from SVI correlate significantly with various societal aspects such as travel patterns, poverty levels, criminal activities, and health behaviors (Fan et al., 2023; Zhang et al., 2019).

Researchers adopt two primary approaches to measure human perception using SVI. The first method involves extracting visual features from SVI to construct an evaluation index based on prior knowledge. This approach effectively assesses residents' evaluations of physical facilities, measuring factors like walkability, rideability, spatial disorder, and impacts on walking behavior in urban environments (Gu et al., 2018; Chen et al., 2023; Jiang et al., 2021).

In contrast, the second approach entails recruiting volunteers to subjectively rate SVI, followed by machine learning techniques to simulate these subjective perceptions. This method aptly gauges residents' "sense of place" in urban areas, based on individual experiences in specific locations (Tuan, 1977). Leveraging SVI and machine learning facilitates extensive perception data collection from volunteers, exploring diverse factors impacting residents' perceptions in urban settings, including safety, aesthetics, psychological stress, urban renewal impacts, playability, and outdoor comfort (Zhang et al., 2018; Han et al., 2022; Ma et al., 2021; Kruse et al., 2021; Liu et al., 2023; Huang et al., 2023). It's important to note that these perceptions are subjective, varying even within the same individual across different scenarios. Tailoring perception data collection to specific study scenarios remains crucial.

Despite the extensive research in measuring subjective perceptions using SVI, there's a dearth of studies evaluating cyclists' subjective perceptions from SVI data. The influence of SVI visual features on cyclists' perceptions of the street environment remains unclear.

2.2. Cyclist's perception of the cycling environment

Cyclists, lacking physical separation from the street environment while riding, possess heightened sensitivity compared to motorized transport users. Studies outlined by Campos Ferreira et al. identify three key dimensions of cyclist perception: Traffic Safety, Public Security, and Comfort/Aesthetic Pleasure (Campos Ferreira et al., 2022).

Regarding Public Security perceptions, cyclists' concerns are influenced by factors like the presence of homeless individuals, deserted areas, poorly lit zones, and gender-related vulnerabilities (Ferrer et al., 2015; Sanders & Judelman, 2018; Russell et al., 2021). Familiarity with routes seems to reduce the impact of public safety concerns on motivation (Ravensbergen et al., 2020).

Studies investigating Traffic Safety perceptions reveal factors such as lane separation, width, colored surfaces, cycling-exclusive zones, and traffic flow significantly influence cyclists' perceptions (von Stülpnagel & Binnig, 2022; Guo et al., 2023; Rivera Olsson & Eeldér, 2023). Positive perceptions align with good cycling facilities and reduced traffic stress (Sanders & Judelman, 2018).

Comfort and Aesthetic Pleasure perceptions among cyclists vary. Some studies find no significant impact on cycling frequency from overhead greenness (Lu et al., 2019), while others note a positive relationship between eye-level greenness and cycling frequency (Gao et al., 2021; Wang et al., 2020).

The prevailing studies on cyclists' perceptions primarily rely on conventional survey-based methods, limiting data volume and cost-effectiveness. Bridging this gap, the combination of street view imagery (SVI) and machine learning presents an opportunity. However, there remains a gap in understanding how visual features from SVI affect cyclists' subjective perceptions. To fill this gap, we propose an SSB (public security, traffic safety, and scenic beauty) framework. By employing this framework, we gathered subjective ratings from regular cyclists, developed models to simulate subjective perceptions, and employed explainable machine learning to discern how visual features influence cyclists' subjective perceptions.

3. Datasets and research methods

3.1. Study area

Shenzhen, a highly developed mega-city situated on the east bank of the Pearl River Estuary adjoining Hong Kong (Fig. 1), forms the focal point of this study. Encompassing Nanshan District, Futian District, and Luohu District, the central urban area spans approximately 344 km² and hosts a resident population of about 4.49 million. Renowned for its robust secondary and tertiary industries, Shenzhen serves as a pivotal center for the science, technology, and headquarters economy sectors.

In recent years, the Shenzhen Municipal Government has dedicated considerable efforts to foster a “bicycle-friendly city.” Their initiatives led to the establishment of approximately 1,759 km of bicycle lanes on both sides, culminating in an average daily ridership of around 850,000 by the close of 2019 (Shenzhen Municipal Bureau of Transportation, 2020). Notably, bicycle transportation has become integral to the daily commutes of residents and functions as a crucial link to public transportation. This study, focusing on Shenzhen, aims to explore the correlation between the city's street environment and cyclists' perceptions of cycling. The findings hold practical significance in advancing cycling infrastructure and advocating sustainable transportation practices.

3.2. Datasets

3.2.1. Street view imagery (SVI)

Researchers can analyze the urban environment from a human perspective by using street view imagery (SVI), which is not possible with aerial or satellite imagery (Biljecki & Ito, 2021). The SVI data used for this work was gathered from Baidu Maps, a highly-ranked electronic map provider in China. To obtain SVI, we selected the midpoint of each road in the network as a sampling point and retrieved the corresponding image from Baidu Maps. At each sampling point, four frames were acquired with the camera facing angles of 0°, 90°, 180°, and 270°. Each frame has a field of view (FOV) of 90°. In total, we acquired 40,776 images from 10,194 points. These images were acquired between 2016 and 2021. Since the study area is situated in the subtropics, the primary vegetation consists of evergreen broadleaf forests, which tend to show minimal seasonal changes. Additionally, given the area's highly

urbanized downtown location, we hypothesized that these sites' landscapes have remained mostly unchanged.

3.2.2. Volunteer rating data

We developed a Street View Rating Software to collect volunteers' subjective perceptions of cycling. Volunteers can simultaneously view a full 360° street environment from one sampling point with four images. Volunteers categorized each scenario as Good or Bad on three dimensions: 'Public Security', 'Traffic Safety', and 'Scenic Beauty'. We invited 50 volunteers, with an equal gender split, to rate 500 scenarios each. This resulted in 50 groups and 75,000 data points.

3.3. Study framework

The study framework, depicted in Fig. 2, begins with employing a fully convolutional network to perform semantic segmentation of street view images, creating a dataset. Volunteers evaluate these images using the Street View Rating Software, aligning their ratings with the segmentation outcomes. This creates volunteer rating datasets reflecting subjective perceptions in three dimensions: public security, traffic safety, and scenic beauty. To address data imbalance, Kmeans-SMOTE processes the volunteer rating datasets for subsequent machine learning. Random forest classifiers are trained for each volunteer's dataset using grid search for optimal hyperparameters, resulting in individual perception models. The volunteers' ratings are quantified to generate perception maps. The SHAP method and Spearman correlation coefficients are combined to analyze how SVI visual features influence cyclists' subjective perceptions.

The subjective perception of volunteers rated SVI within the SSB framework, emphasizing three key dimensions: public security, traffic safety, and scenic beauty. Inspired by Campos Ferreira et al. (2022), this framework prioritizes cyclists' visual subjective perception. Traffic safety relates to transportation visibility and cyclist confidence in street safety. Public security involves perceptions of crime risk while cycling. Scenic beauty focuses on the aesthetic pleasure of the surroundings.

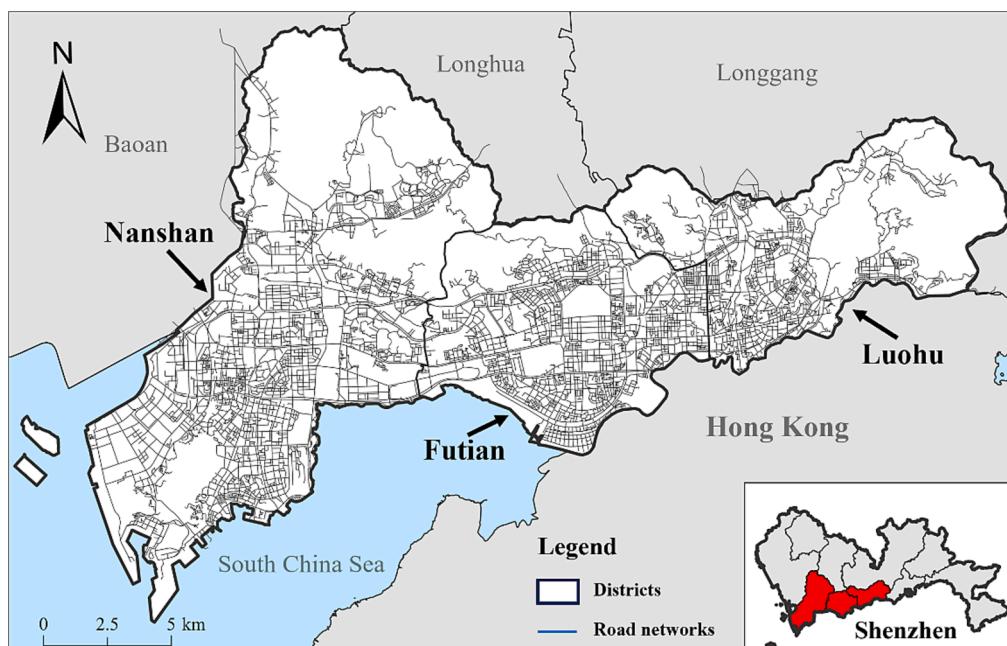


Fig. 1. Study area.

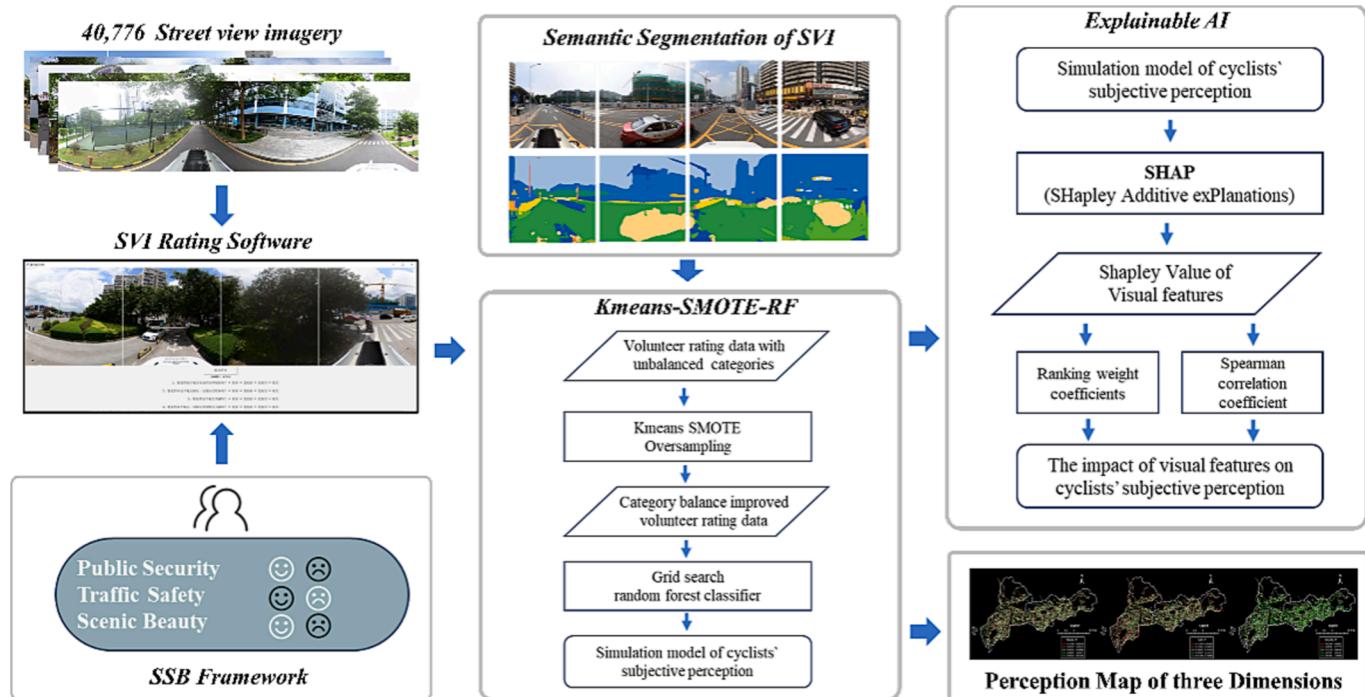


Fig. 2. Study Framework.

3.4. Methodology

3.4.1. Semantic segmentation of SVI based on fully convolutional networks

Semantic segmentation is integral to our approach, examining how visual components influence a cyclist's perception. This process entails dividing an image into semantically meaningful regions, calculating pixel percentages under category masks, and downsizing the image into a one-dimensional array for machine learning.

In this study, we employed a fully convolutional neural network developed by Yao et al. for semantic segmentation. This network was trained on the ADE-20 K dataset, achieving an accuracy of 81.44 % on the training set and 66.83 % on the validation set (Yao et al., 2019). After the street view data segmentation, we discovered that 19 categories of features made up 96.1927 % of the visual features, while the remaining 131 categories only accounted for 3.8073 % (Fig. 3). To simplify the analysis, we focused solely on these 19 categories of visual

features.

3.4.2. KmeansSMOTE RF model

The volunteer rating data often exhibits category imbalance, impacting the model's ability to differentiate positive and negative examples. To address this, the Kmeans SMOTE (Kmeans Synthetic Minority Over-Sampling Technique) algorithm tackles sample class imbalance effectively (Douzas et al., 2018). In contrast to traditional SMOTE methods, Kmeans SMOTE utilizes K-means clustering to identify regions in the data space for oversampling. By determining the number of samples to generate based on the density of each data space, Kmeans SMOTE effectively mitigates the risk of generating noisy data and excels in balancing classes. This approach ensures that the oversampled data closely aligns with the original dataset, preventing excessive disparity between the two (Fonseca et al., 2021). This method comprises three key steps: clustering, filtering, and oversampling. Initially, the input space is

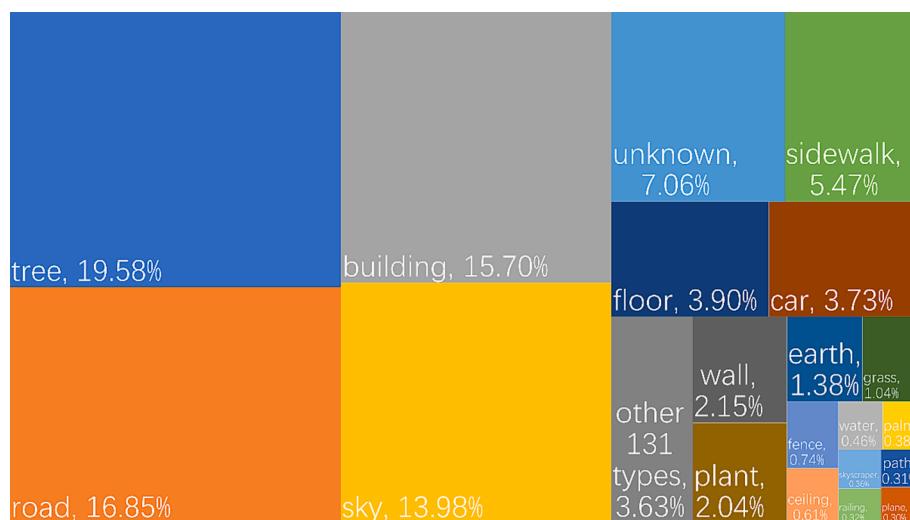


Fig. 3. Visual Features Percentage Diagram.

divided into k groups based on Imbalance Ratio (IR) using K-means clustering. High-proportion minority class clusters are identified in the filtering step for oversampling. Synthetic sample generation is determined by allocating more samples to clusters with fewer minority instances, followed by SMOTE application for achieving the desired minority-majority ratio.

In this study, 80 % of volunteer rating data was allocated to a training set and 20 % to a validation set. K-means with a 0.05 cluster balance threshold was applied to the training set, employing Kmeans-SMOTE on each volunteer's rating data. The resultant data underwent individual training using a random forest classifier, and hyperparameters were optimized through grid search in each training session. This approach balances category imbalances in volunteer rating data, enhancing the model's ability to discern positive and negative instances.

3.4.3. Precision validation

The study employed Area Under the Curve (AUC) and Accuracy Score to assess the model's predictive accuracy. AUC is a metric gauging binary classification model effectiveness via the Receiver Operating Characteristic (ROC) curve. This curve illustrates the relationship between the true positive rate (TPR) and false positive rate (FPR) at different thresholds. AUC values range from 0.5 to 1, where 0.5 signifies performance akin to random guessing, and 1 indicates accurate prediction of both positive and negative instances.

Accuracy Score measures the proportion of correctly classified samples out of the total. The confusion matrix depicts True Positive (TP) for accurately predicted positives, True Negative (TN) for accurately predicted negatives, False Positive (FP) for wrongly predicted positives, and False Negative (FN) for wrongly predicted negatives. Accuracy Score is calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Where *Accuracy* denotes the accuracy of the model; *TP*, *TN*, *FP*, and *FN* represent the four categorization cases of the test sample, as defined above.

3.4.4. Quantification of subjective perception

Following Kmeans-SMOTE-RF training, we create models that match the number of volunteers. These models can imitate the perceptual ratings of the volunteers. We then used these models to generate ratings for all samples and determine the ratings of the volunteer population for each sampling point using a rating ratio. The positive ratio and negative ratio were defined as follows:

$$Pi = \frac{pi}{pi + ni} \quad (2)$$

$$Ni = \frac{ni}{pi + ni} \quad (3)$$

Where *Pi* and *Ni* denote the proportion of image *i* that is evaluated positively and negatively, respectively. *pi* is the number of times image *i* is evaluated as positive and *ni* is the number of times image *i* is evaluated as negative.

3.4.5. Interpretation of models via SHAP

SHAP (Shapley Additive Explanation) is an explanatory model rooted in Shapley values, a concept from cooperative game theory (Barredo Arrieta et al., 2020; Lundberg & Lee, 2017). In SHAP, features within a model act as cooperative game participants generating collective benefits. Each feature contributes differently, positively or negatively, to the model's predictions. The Shapley value quantifies these contributions, allowing us to assess individual elements' positive or negative impact on the sample and determine their influence on the model's predictions.

The Shapley value is calculated as follows:

$$\phi_i(f, x) = \sum_{R \in \mathcal{R}} \frac{1}{M!} [f_x(P_i^R \cup \{i\}) - f_x(P_i^R)] \quad (4)$$

Where $\phi_i(f, x)$ represents the influence of feature *x* on the outcome predicted by model *f*, i.e., feature *x*'s Shapley value. \mathcal{R} represents the set of all feature arrangements, *M* is the number of input features in the model, P_i^R denotes the set of all features before feature *i*, and $f_x(P_i^R \cup \{i\}) - f_x(P_i^R)$ represents the contribution generated by feature *i* after it is added.

As a cooperative game-theoretic solution concept, the Shapley value of all features in one model also obeys the following equation:

$$y_i = y_{base} + f(x_{i1}) + f(x_{i2}) + \dots + f(x_{ij}) \quad (5)$$

where y_i represents the predicted value of the model for sample *i*, y_{base} represents the baseline value of the whole model, and $f(x_{ij})$ is the Shapley value of the characteristic feature *j* of sample *i*. When $f(x_{ij}) > 0$, it represents that feature *j* has a positive effect; when $f(x_{ij}) < 0$, it represents that feature *j* has a negative effect. And the larger the absolute value of $f(x_{ij})$ is, the stronger the influence of feature *j* is.

In Equation (5), when there's significant variation in the model's feature sensitivity, averaging operations in SHAP across multiple models might drive the Shapley value closer to 0. This limitation hampers our ability to discern changes in each feature's value, the extent of positive and negative impact, and the magnitude of alterations in the model's predictions. Calculating the Shapley value for positive cycling perception prediction, we then measured Spearman correlations between visual feature values and Shapley values. This analysis reveals which visual features positively influence a cyclist's perception. An increase in a visual feature's value, indicated by a positive Spearman correlation coefficient, is likely to enhance a cyclist's perception of riding. Conversely, a negative correlation signifies that an increase in the visual feature's value might negatively impact the cyclist's perception.

The Shapley value gauges the strength of visual features in impacting the model's predictions. However, the Spearman correlation coefficient solely signifies the monotonic relationship between variables, lacking the ability to portray their absolute values' magnitude. Therefore, an additional parameter is required to assess each visual feature's influence strength. While the Shapley value can be positive or negative, indicating the feature's impact on prediction probabilities, directly correlating the Spearman coefficient with the Shapley value might complicate analysis due to strong influences at both ends of the axis. To address this, we devised ranking weight coefficients to measure each visual feature's influence using the following formula:

$$Weight = 1 - (rank * 0.05) \quad (6)$$

where *Weight* is the ranking weight coefficient and *rank* is the ascending ranking place of the absolute value of the Shapley value. The weights represent the intensity of the contribution of a visual element to influence volunteers to make the corresponding evaluation.

4. Results and discussion

4.1. Comparison of different oversampling and Machine learning methods

This paper introduces a comprehensive machine learning framework that comprises an oversampling algorithm and a classifier to address data imbalance in volunteer scoring data and model the scoring behavior of individual volunteers. The oversampling algorithm aims to mitigate class imbalance, while the classifier captures the nuanced perception of cyclists regarding the street environment. Robust oversampling algorithms and classifiers serve as a foundation for subsequent explanatory machine learning, elucidating the impact of street environment factors on subjective cyclist perception. The comparative evaluation of oversampling algorithms and classifiers across five key

parameters, including accuracy, AUC, kappa coefficient, errors of commission, and errors of omission, ensures a thorough assessment of their effectiveness and suitability for the proposed framework.

In our study, we compared four oversampling algorithms: Original SMOTE, Borderline SMOTE, SVM SMOTE, and Kmeans SMOTE. We controlled for the variable of the classifier and evaluated each of the four oversampling algorithms after integrating them with a random forest classifier to isolate the impact of oversampling techniques on model performance while maintaining consistency in classifier methodology across comparisons.

Table 1 lists the average accuracy results of the four oversampling algorithms with the Random Forest classifier on 50 volunteer rating data, on three perceptual dimensions, including public security, traffic safety, and scenic beauty. Notably, Kmeans SMOTE consistently exhibits superior performance with the highest average AUC, average precision, and average Kappa coefficient across all dimensions. While it may not top the charts in terms of errors of commission and omission, Kmeans SMOTE still demonstrates commendable performance in these aspects. Therefore, Kmeans SMOTE oversampling algorithm is selected in this paper.

In the Machine Learning Classifiers section, we compare the performance of Random Forest (RF), Support Vector Machine (SVM), eXtreme Gradient Boosting (Xgboost), and Decision Tree (DT) classifiers in conjunction with the KmeansSMOTE oversampling algorithm on the volunteer rating data. By comparing the classifiers with this oversampling technique, we aim to determine the most effective approach for addressing class imbalance in the dataset.

Table 2 lists the average accuracy results of the four classifiers with the Kmeans SMOTE oversampling algorithm on 50 volunteer rating data, on three perceptual dimensions. Among the classifiers, Random Forest (RF) stands out with the best average AUC, average precision, average Kappa coefficient, and average errors of commission across all dimensions. While its performance in average errors of omission may not rank the highest, RF still maintains a commendable level.

In summary, this paper adopts the Kmeans SMOTE oversampling algorithm in conjunction with the Random Forest (RF) classifier, forming the Kmeans SMOTE RF machine learning approach.

4.2. Performance of the KmeansSMOTE-RF model in volunteer rating data with class imbalance

The performance evaluation of the KmeansSMOTE-RF model was carried out concerning class imbalance in volunteer rating data. The imbalance ratio (IR) was computed by dividing negative samples by positive samples, illustrating imbalances in the dataset's class distribution. Probability density plots (**Fig. 4**) visualized the imbalance ratios across the three perceptual dimensions. Notably, volunteer rating data displayed varying levels of imbalance across the public security, traffic safety, and scenic beauty dimensions.

An imbalance ratio of less than 1 indicates that there are more positive than negative samples in this volunteer's rating data. Additionally,

the kurtosis, which characterizes the dataset's distribution, reveals that the volunteer rating dataset are predominantly centered around the median imbalance ratio (IR). For instance, the public security dimension exhibited the highest imbalance with a median IR of 0.275 and kurtosis of 16.444, indicating a substantial imbalance favoring positive samples. Similarly, the traffic safety and scenic beauty dimensions showed imbalances, with median IRs of 0.883 and 0.755, respectively. Imbalanced learning, a prevalent issue when training models with uneven class distributions, was noted in this context. For instance, the public security dimension exhibited the highest imbalance with a median IR of 0.275 and kurtosis of 16.444, indicating a substantial imbalance favoring positive samples. Similarly, the traffic safety and scenic beauty dimensions showed imbalances, with median IRs of 0.883 and 0.755, respectively. Imbalanced learning, a prevalent issue when training models with uneven class distributions, was noted in this context.

Addressing imbalanced class distributions during model training poses a significant challenge known as the imbalanced learning problem ([Chawla et al., 2004](#)). This issue arises in binary classification scenarios where data distribution across different classes is uneven ([Abdi & Hashemi, 2016](#)). Models trained on such imbalanced data tend to favor majority classes while neglecting minority ones. Imbalanced learning is notably encountered in land-use and land-cover classification (LULC classification) within geographical research. The skewed distribution of land-use types presents a substantial hurdle for scholars. To combat this challenge, researchers have explored methods like KmeansSMOTE, which has demonstrated improvements in outcomes ([Fonseca et al., 2021](#)).

In evaluating the performance of the KmeansSMOTE-RF model on imbalanced volunteer rating data, we compared models trained with and without oversampling using KmeansSMOTE. The model evaluation was based on accuracy and AUC metrics, where grid search was utilized to optimize hyperparameters for each model.

Initially, models trained without KmeansSMOTE oversampling exhibited high accuracy, but the distribution of AUC was close to 0.5, indicating classification performance akin to random guessing (**Fig. 5a** and **5b**). Conversely, models trained post KmeansSMOTE oversampling (**Fig. 5c** and **5d**) displayed a notable improvement in AUC values. The majority of models showcased higher AUC values, particularly in the public security dimension, with a peak at 0.86, signifying improved discrimination between positive and negative examples. Despite less concentration in AUC values for the scenic beauty and traffic safety dimensions, most models still achieved an AUC greater than 0.60.

Comparing accuracy metrics between models trained with and without KmeansSMOTE oversampling (**Fig. 5a** vs. **5c**), models benefiting from KmeansSMOTE exhibited superior performance. Most models achieved an accuracy above 0.70 in public security and scenic beauty dimensions, while the traffic safety dimension showed slightly lower but still favorable accuracy, mostly above 0.60.

Fig. 6 indicates a significant enhancement in AUC and accuracy metrics when employing KmeansSMOTE for data processing. The models for public security experienced an average AUC enhancement of

Table 1

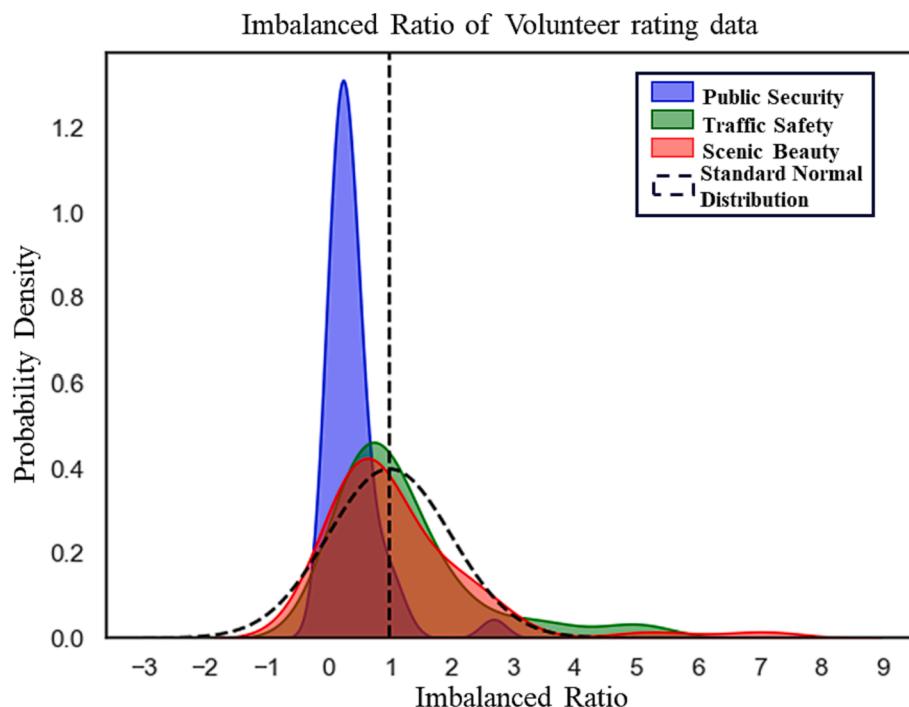
Accuracy result of four oversampling algorithms on three dimensions of cycling volunteer rating data.

Dimension	Oversampling Algorithms	AUC	Accuracy	Kappa coefficient	Errors of commission	Errors of omission
Public Security	Kmeans-SMOTE RF	0.824	0.824	0.648	0.247	0.105
	Original SMOTE RF	0.807	0.805	0.611	0.168	0.219
	SVM-SMOTE RF	0.795	0.809	0.599	0.277	0.133
	Borderline-SMOTE RF	0.797	0.796	0.592	0.186	0.219
Traffic Safety	Kmeans-SMOTE RF	0.700	0.698	0.398	0.305	0.296
	Original SMOTE RF	0.686	0.683	0.369	0.299	0.329
	SVM-SMOTE RF	0.682	0.685	0.366	0.311	0.326
	Borderline-SMOTE RF	0.684	0.682	0.366	0.308	0.324
Scenic Beauty	Kmeans-SMOTE RF	0.715	0.715	0.429	0.312	0.259
	Original SMOTE RF	0.702	0.700	0.401	0.281	0.314
	SVM-SMOTE RF	0.690	0.695	0.382	0.315	0.305
	Borderline-SMOTE RF	0.700	0.696	0.397	0.282	0.318

Table 2

Accuracy result of four classifiers on three dimensions of cycling volunteer rating data.

Dimension	Classifier	AUC	Accuracy	Kappa coefficient	Errors of commission	Errors of omission
Public Security	RF	0.824	0.824	0.648	0.247	0.105
	Xgboost	0.805	0.805	0.609	0.218	0.172
	SVM	0.763	0.763	0.526	0.279	0.195
	DT	0.745	0.745	0.490	0.248	0.262
Traffic Safety	RF	0.700	0.698	0.398	0.305	0.296
	Xgboost	0.692	0.691	0.382	0.312	0.304
	SVM	0.655	0.651	0.310	0.331	0.358
	DT	0.646	0.643	0.290	0.325	0.382
Scenic Beauty	RF	0.715	0.715	0.429	0.312	0.259
	Xgboost	0.699	0.699	0.397	0.310	0.292
	SVM	0.666	0.663	0.330	0.327	0.342
	DT	0.663	0.663	0.325	0.321	0.352

**Fig. 4.** Probability density plots of imbalance ratios for samples of the three perceptual dimensions.

0.327, while traffic safety and scenic beauty models had an average AUC enhancement of 0.200 and 0.209, respectively. Although a slight decrease (-0.0014) in accuracy was observed in the public security model, there was an average enhancement of 0.171 for the traffic safety model and 0.156 for the scenic beauty model, signifying the overall improvement in model performance.

4.3. Spatial pattern of cyclists' subjective perception

Utilizing the KmeansSMOTE-RF model, we simulated cycling perception across 10,194 scenes for each volunteer. These simulations enabled the evaluation of street cycling environments based on positive-to-negative evaluation ratios, which we visualized for the three dimensions. In terms of public security, Nanshan, Futian, and Luohu, the well-developed urban districts in Shenzhen, generally exhibit higher perceived public security levels (Fig. 7a). However, some secondary roads in the southern part of Nanshan display lower perceived public security. In relation to traffic safety, main roads in the network center portray higher safety perceptions, while secondary roads on the urban periphery showcase lower perceived safety levels (Fig. 7b). For scenic beauty, areas with high perception levels intermingle with regions of lower perception (Fig. 7c).

To better understand the spatial distribution of cyclists' perceptions, we employed the Getis-Ord Gi^* statistic to analyze cyclists' perceptions spatially. Fig. 8 displays the clustering patterns across three dimensions, while Fig. 9 illustrates typical scenarios within high and low-value clustering areas for each dimension.

In terms of public security perceptions, Shenzhen's primary urban areas exhibit both high and low-value clusters. High-value clusters are notably located in Luohu's western areas and central and eastern Futian, whereas low-value clusters are concentrated in Nanshan's southwestern and northern parts, with scattered distribution in Luohu's periphery (Fig. 8a). Correspondingly, high-value areas typically manifest well-constructed, green spaces, indicating higher management levels, while low-value areas are often undeveloped or characterized by messy building facades (Fig. 9).

For cyclists' traffic safety perception, Shenzhen's main urban regions portray moderate scores, with high-value clusters concentrated in central districts and low-value clusters at the edges. Predominant high-value areas are observed in central Futian, southwestern Luohu near Futian, and the center of Nanshan. Conversely, low-value clusters mostly concentrate in Nanshan's southwestern and northern sectors, along with sporadic occurrences in Luohu's periphery (Fig. 8b). Fig. 9 shows that higher traffic safety perception aligns with well-designed roads with

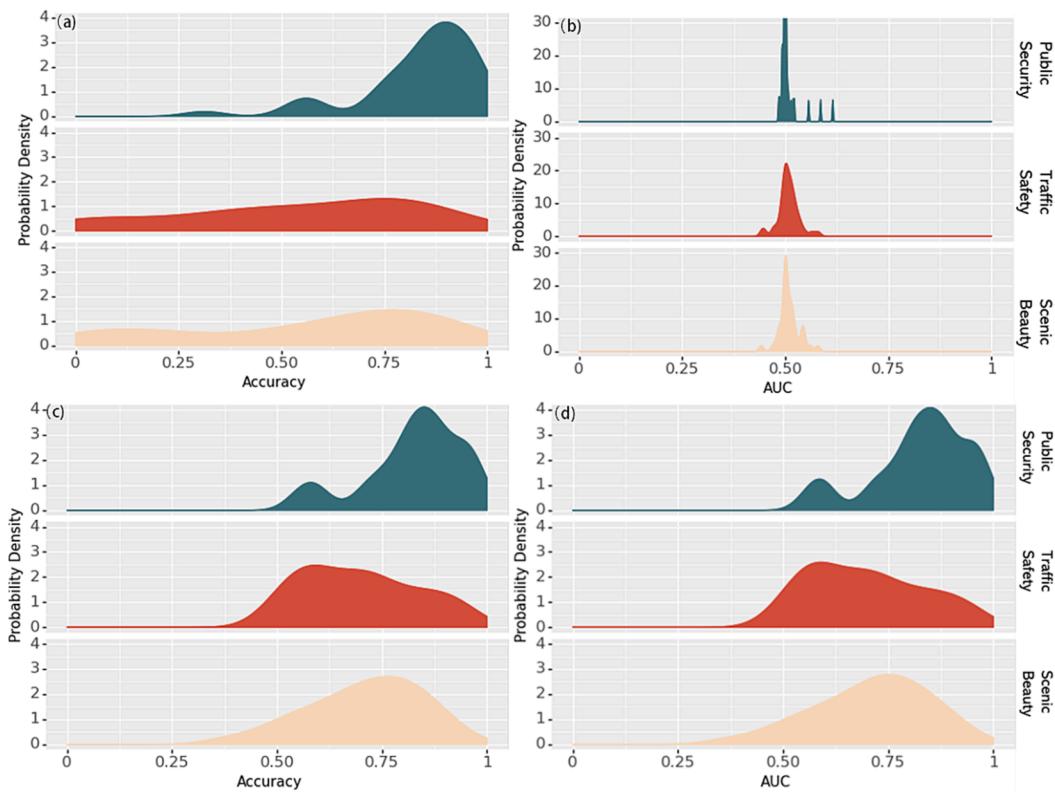


Fig. 5. Probability density plots of the accuracy and AUC of the model before and after KmeansSMOTE oversampling: (a) accuracy of models trained on data before KmeansSMOTE oversampling; (b) AUC of models trained on data before KmeansSMOTE oversampling ; (c) accuracy of models trained on data after KmeansSMOTE oversampling; (d) AUC of models trained on data after KmeansSMOTE oversampling.

segregated lanes and poor lighting correlates with lower cyclist safety perception.

Cyclists' perception of scenic beauty in Shenzhen's urban zones averages a median score of 0.56, with clusters of high and low values being small and dispersed. High-value clusters scatter in Luohu's northern and western regions, Futian's northern, southern, and western parts, and central and southern Nanshan, primarily at the districts' edges. Low-value clusters concentrate in Nanshan's southwestern and northern parts and Futian's southern sector (Fig. 8c). Fig. 9 reveals that high-value areas tend to be greener with well-maintained streets, while low-value areas lack greenery, feature temporary walls, or experience heavy traffic.

4.4. The impact of visual features on cyclists' subjective perception

In Equation (5), averaging SHAP values from models with varying feature sensitivities results in convergence to 0. To overcome this, we integrated Spearman correlation coefficients and ranking weighting coefficients to characterize visual elements' effects on ride perception across multiple models. The four-quadrant plot showcases the relationship:

The vertical axis displays Spearman correlation coefficients between visual feature values and SHAP values. Higher positive correlation coefficients indicate an increased likelihood of a positive classification with rising feature values. Conversely, higher negative correlation coefficients indicate a greater chance of a negative classification with increased feature values. The horizontal axis represents the ranking weight coefficient. Higher values signify stronger visual feature influences on the final judgment result of a sample. In the P-H quadrant, visual features suggest an increased likelihood of positive classification with a more pronounced influence. Conversely, in the N-H quadrant, visual features indicate a higher chance of negative classification with a more substantial impact.

Fig. 10a exhibits a four-quadrant plot illustrating the influence of visual features on cyclists' perception of public security. Notably, the P-H quadrant encompasses trees and roads, signifying that expansive and verdant roadways tend to evoke higher perceptions of public security. Conversely, visual elements in the N-H quadrant such as sky, grass, palm trees, walls, bare soil, and unidentified features suggest lower security perceptions, indicating cluttered or underdeveloped environments.

In **Fig. 10b**, the impact of various visual features on cyclists' perception of traffic safety is depicted. Visual elements in the P-H quadrant, including fences, sidewalks, trees, bare soil, roads, and sky, contribute to an enhanced perception of traffic safety. In contrast, the N-H quadrant features like buildings, unrecognizable elements, and rails correspond to decreased traffic safety perception, particularly in narrow, congested areas.

Notably, among the visual features in **Fig. 10b**, the highest positive correlation is observed between fences and sidewalks. This association underscores that roads equipped with these facilities offer segregated and spacious paths for cyclists, which aligns with findings from Nolan et al. (Nolan et al., 2021). Their study highlights that separating motor vehicles from cyclists through designated lanes or railings significantly reduces cyclist stress and enhances their sense of safety during transit.

Fig. 10c explores how visual features influence cyclists' perception of scenic beauty. The P-H quadrant features roads and trees, indicating that wider roads and abundant greenery contribute to a more aesthetically pleasing cycling environment. Conversely, elements in the N-H quadrant, such as sky, unrecognizable features, water, palm trees, fences, cars, and walls, imply that cluttered, narrow, or busy streets may diminish the perceived beauty of the surroundings for cyclists.

4.5. The impact of architectural features and natural features on the subjective perception of cyclists

Examining the amalgamated influence of architectural and natural

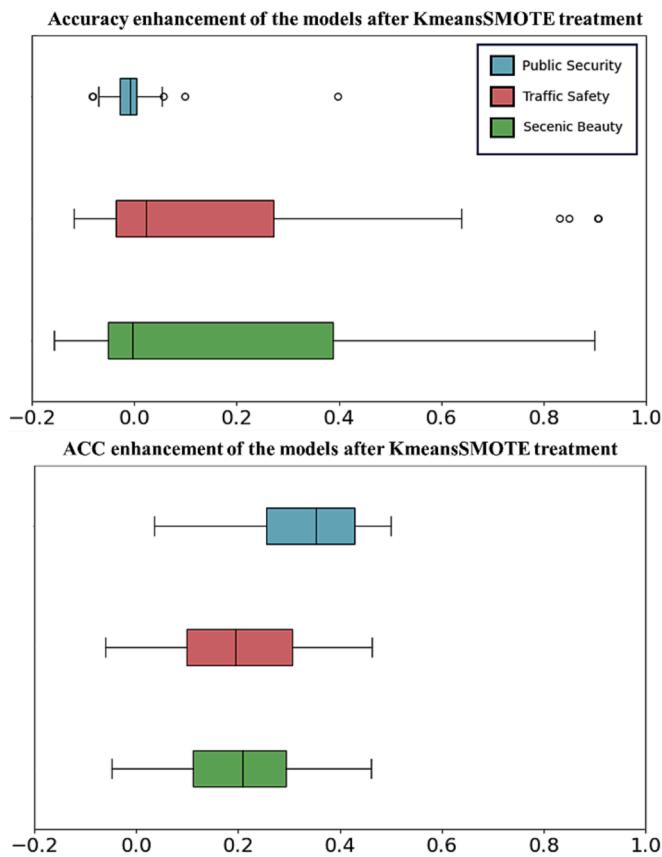


Fig. 6. Enhancement in accuracy and AUC of models using training data processed by KmeansSMOTE.

features on cyclists' subjective perception across three dimensions unveils notable insights.

Architectural features encompass roads, ceilings, floors, skyscrapers, and walls, significantly impacting cyclists' perceptions negatively in various dimensions. Walls, known for obstructing views and sunlight, often lead to a "wall effect," inducing frustration, boredom, and feelings of insecurity (Wong et al., 2011; Zhang et al.). Similarly, skyscrapers contribute to this effect, while enclosed spaces created by ceilings and paved floors may evoke tunnel-like environments, unfavorably shaping cyclists' perceptions. Thus, contemporary urban planning advocates considering the impact of these architectural elements on cyclists' perceptions of public security and scenic beauty. Regarding roads, their width affects greenery perception, with peaks observed at widths of 35 m and 50 m (Ma et al., 2021). Our study noted a positive correlation between road image pixels and public security and scenic beauty perceptions among cyclists. However, the influence on traffic safety lacks a clear correlation with the SHAP value, requiring further investigation.

Natural features, such as trees, bare soil, the sky, and palm trees, contribute distinctively to cyclists' environmental perception. Trees elicit positive responses across all dimensions, promoting tranquility and aligning with streetscape art theories (Ashihara, 1983; Kaplan & Kaplan, 1989). Empirical studies endorse green surroundings as conducive to cycling (Gao et al., 2021). However, palm trees negatively impact perception in all dimensions. While the sky generally uplifts urban environments (Asgarzadeh et al., 2014), its effect on cycling perception varies. Open skies discourage weekday cycling but encourage weekend cycling (Bai et al., 2023). Our findings highlight the sky's fluctuating impact on cyclists' safety and scenery perceptions, particularly in traffic safety perception, necessitating further investigation.

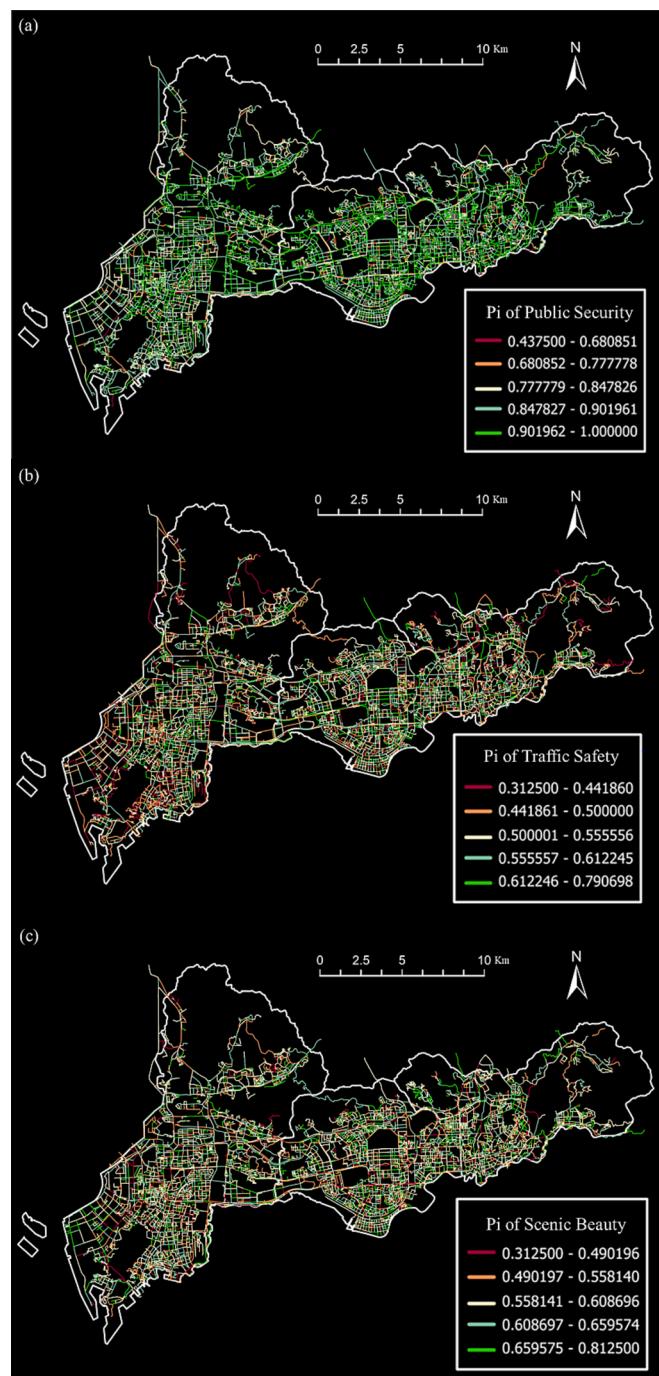


Fig. 7. Perception maps of cyclists' cycling. (a) perception map of the public security dimension; (b) perception map of the traffic safety dimension;(c) perception map of the scenic beauty dimension.

5. Conclusions

Subjective perceptions of street suitability for cycling can impact cyclists' willingness to ride. The use of SVI to simulate the subjective visual perception of cyclists as they ride can help optimize the cycling environment in urban areas. In this study, we propose a SSB framework (Public Security, Traffic Safety, Scenic Beauty), and quantify cyclists' subjective perception form SVI and volunteer rating data. The major contribution and findings of this study can be summarized as follows.

Firstly, applying KmeansSMOTE to process volunteer rating data while modeling cyclists' subjective perceptions with a Random Forest classifier increased AUC by 0.327 in public security, 0.200 in traffic

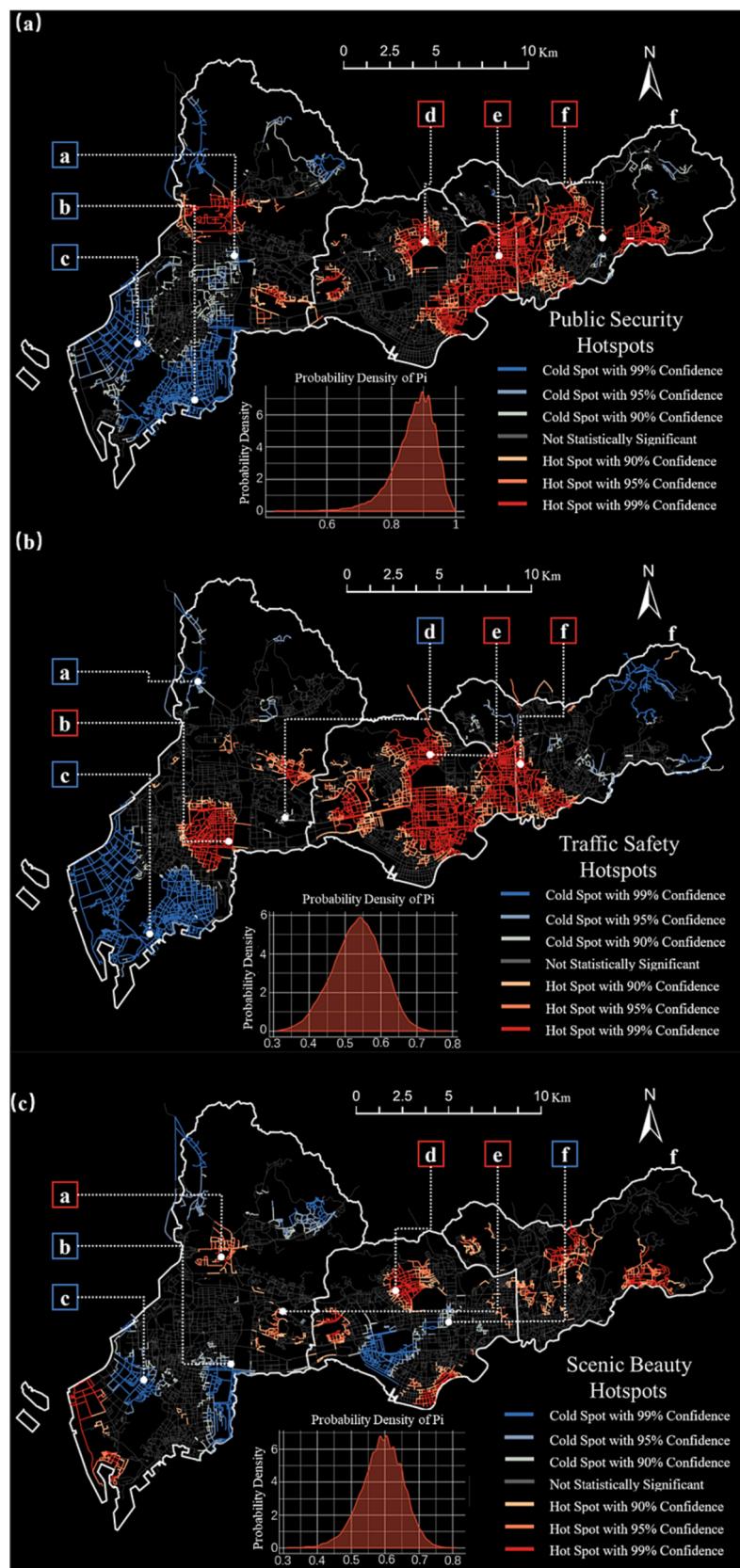


Fig. 8. Spatial clustering maps of cyclists' subjective perceptions: (a) Spatial clustering of cyclists' subjective perceptions of public security dimension;(b) Spatial clustering of cyclists' subjective perceptions of traffic safety dimension; (c) Spatial clustering of cyclists' subjective perceptions of scenic beauty dimension.



Fig. 9. Typical Scenarios of High and Low Subjective Perceived High Value and Low Value for Cyclists in Three Dimensions of Public Security, Traffic Safety, and Scenario Beauty. The left column is the high-value scenario and the right column is the low-value scenario.

safety, and 0.209 in scenic beauty. This highlights the effectiveness of KmeansSMOTE in enhancing model discrimination and mitigating overfitting issues arising from limited and imbalanced volunteer data. Residents' concerns about subjective perception dimensions vary across different scenarios. Collecting scenario-specific subjective perception data is essential but challenging due to issues related to small and uneven sample sizes. In this context, KmeansSMOTE-RF proves valuable, significantly improving model accuracy and AUC. This technique aids researchers in understanding residents' subjective perceptions in specific scenarios and deciphering the complex ways in which urban characteristics influence them.

Secondly, we investigate how the visual features of SVI impact the subjective perceptions of cyclists, with the combination of SHAP and Spearman correlation. We found that Tree exhibited a notably positive influence on all three dimensions of cyclists' perceptions, whereas the presence of palm trees had an adverse effect on cyclists' subjective perceptions of traffic safety. Additionally, walls consistently yielded negative outcomes, consistent with previous research. In contrast, both fences and sidewalks emerged as strong contributors to enhancing cyclists' subjective perceptions of traffic safety. These findings offer valuable insights for the design of more cyclist-friendly urban environments and strategies to encourage greater cycling participation among

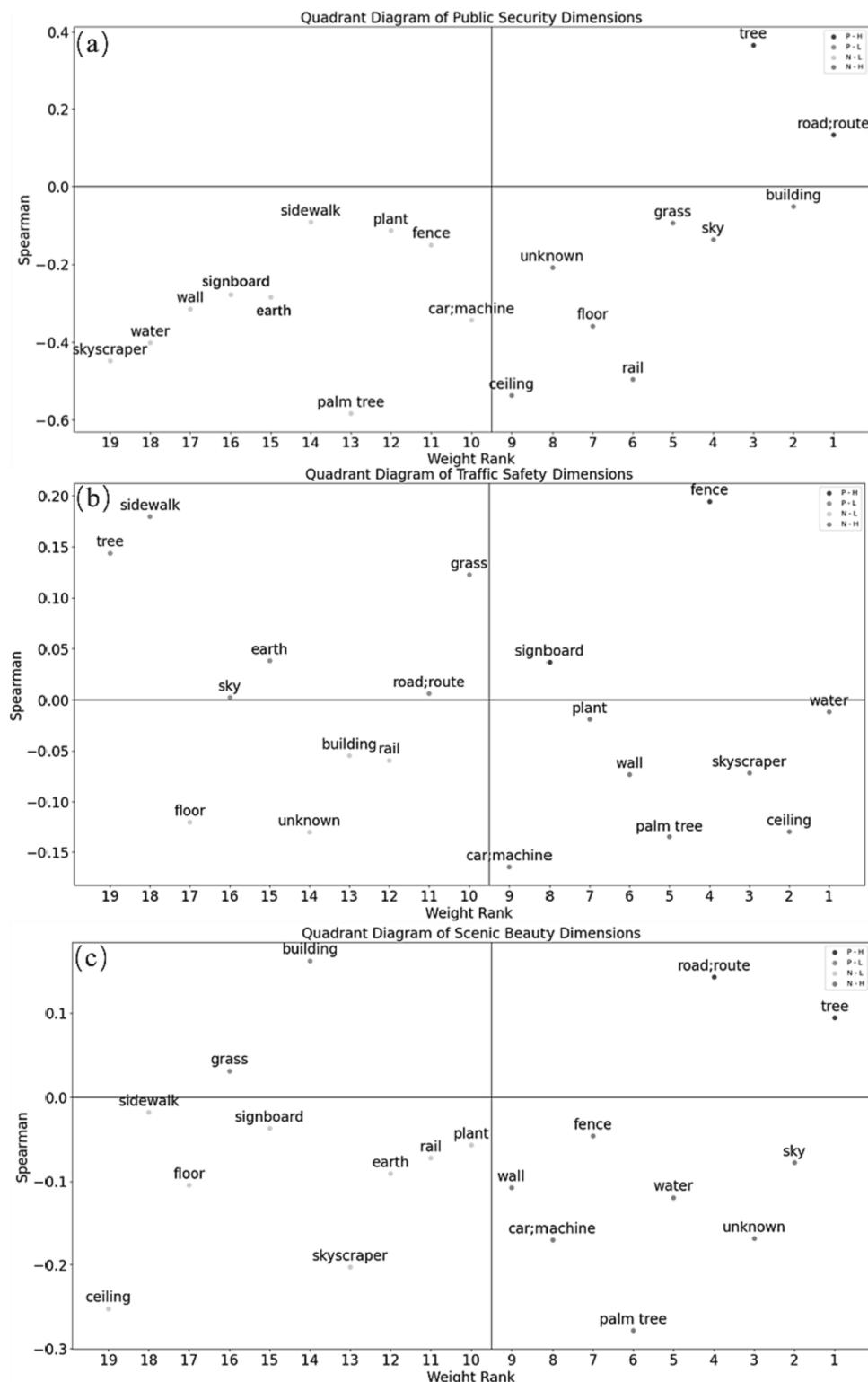


Fig. 10. Four-quadrant plot of the effect of visual features on cyclists' subjective perceptions of three dimensions: (a)Quadrant Diagram of Public Security Dimensions; (b) Quadrant Diagram of Traffic Safety Dimensions; (c)Quadrant Diagram of Scenic Beauty Dimensions.

residents.

Despite the merit of this study, we also have to discuss the limitations of this study and directions for future improvement efforts. First, the idea of the modeling in the present study is to train a large number of models to 'customize' the model of each volunteer to simulate his or her perception. The disadvantage of this is that some of the volunteers were not modeled very well. Regarding the Kmeans-SMOTE-RF model, not all

volunteer rating data could be successfully clustered, and the rating data of four volunteers could not be clustered because the categories were too unbalanced, so they were temporarily excluded from the dataset of the present study. Another aspect that needs more attention is that it remains unclear whether gender play a role in shaping perceptions. Our future work will further investigate the gender differences in the impact of visual features on cyclists' subjective perception.

- Montello, D.R., Goodchild, M.F., Gottsegen, J., Fohl, P., 2003. Where's downtown?: behavioral methods for determining references of vague spatial queries. *Spat. Cogn. Comput.* 3 (2–3), 185–204. <https://doi.org/10.1080/13875868.2003.9683761>.
- Nolan, J., Sinclair, J., Savage, J., 2021. Are bicycle lanes effective? the relationship between passing distance and road characteristics. *Accid. Anal. Prev.* 159, 106184 <https://doi.org/10.1016/j.aap.2021.106184>.
- Ravensbergen, L., Buliung, R., Laliberté, N., 2020. Fear of cycling: social, spatial, and temporal dimensions. *J. Transp. Geogr.* 87, 102813 <https://doi.org/10.1016/j.jtrangeo.2020.102813>.
- Rivera Olsson, S., Elldét, E., 2023. Are bicycle streets cyclist-friendly? micro-environmental factors for improving perceived safety when cycling in mixed traffic. *Accid. Anal. Prev.* 184, 107007 <https://doi.org/10.1016/j.aap.2023.107007>.
- Russell, M., Davies, C., Wild, K., Shaw, C., 2021. Pedalling towards equity: exploring women's cycling in a New Zealand city. *J. Transp. Geogr.* 91, 102987 <https://doi.org/10.1016/j.jtrangeo.2021.102987>.
- Sanders, R.L., Judelman, B., 2018. Perceived safety and separated bike lanes in the Midwest: results from a roadway design survey in Michigan. *Transp. Res. Rec.* 2672 (36), 1–11. <https://doi.org/10.1177/0361198118758395>.
- Shen, Y., Zhang, X., Zhao, J., 2018. Understanding the usage of dockless bike sharing in Singapore. *Int. J. Sustain. Transp.* 12 (9), 686–700. <https://doi.org/10.1080/15568318.2018.1429696>.
- Shenzhen Municipal Bureau of Transportation. (2020). Shenzhen City Bicycle Traffic Development Plan(2021–2035). Retrieved from <http://jtzs.sz.gov.cn>.
- Tu, Y., Chen, P., Gao, X., Yang, J., Chen, X., 2019. How to make dockless bikeshare good for cities: curbing oversupplied bikes. *Transp. Res. Rec.* 2673 (6), 618–627. <https://doi.org/10.1177/0361198119837963>.
- Tuan, Y.-F., 1977. *Space and place: the perspective of experience*. U of Minnesota Press.
- von Stülpnagel, R., Binnig, N., 2022. How safe do you feel? – a large-scale survey concerning the subjective safety associated with different kinds of cycling lanes. *Accid. Anal. Prev.* 167, 106577 <https://doi.org/10.1016/j.aap.2022.106577>.
- Wang, R., Lu, Y., Wu, X., Liu, Y., Yao, Y., 2020. Relationship between eye-level greenness and cycling frequency around metro stations in Shenzhen, China: a big data approach. *Sustain. Cities Soc.* 59, 102201 <https://doi.org/10.1016/j.scs.2020.102201>.
- Winters, M., Brauer, M., Setton, E.M., Teschke, K., 2013. Mapping bikeability: a spatial tool to support sustainable travel. *Environ. Plann. b. Plann. Des.* 40 (5), 865–883. <https://doi.org/10.1068/b38185>.
- Winters, M., Teschke, K., 2010. Route preferences among adults in the near market for bicycling: findings of the cycling in cities study. *Am. J. Health Promot.* 25 (1), 40–47. <https://doi.org/10.4278/ajhp.081006-QUAN-236>.
- Wong, M.S., Nichol, J., Ng, E., 2011. A study of the “wall effect” caused by proliferation of high-rise buildings using GIS techniques. *Landsl. Urban Plan.* 102 (4), 245–253. <https://doi.org/10.1016/j.landurbplan.2011.05.003>.
- Xu, Y., Chen, D., Zhang, X., Tu, W., Chen, Y., Shen, Y., Ratti, C., 2019. Unravel the landscape and pulses of cycling activities from a dockless bike-sharing system. *Comput. Environ. Urban Syst.* 75, 184–203. <https://doi.org/10.1016/j.compenvurbsys.2019.02.002>.
- Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., Wang, R., Wang, J., Guan, Q., 2019. A human-machine adversarial scoring framework for urban perception assessment using street-view images. *Int. J. Geogr. Inf. Sci.* 33 (12), 2363–2384. <https://doi.org/10.1080/13658816.2019.1643024>.
- Zhang, R., Hanaoka, T., 2022. Cross-cutting scenarios and strategies for designing decarbonization pathways in the transport sector toward carbon neutrality. *Nat. Commun.* 13 (1), 3629. <https://doi.org/10.1038/s41467-022-31354-9>.
- Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., Ratti, C., 2018. Measuring human perceptions of a large-scale urban region using machine learning. *Landsl. Urban Plan.* 180, 148–160. <https://doi.org/10.1016/j.landurbplan.2018.08.020>.
- Zhang, F., Wu, L., Zhu, D., Liu, Y., 2019. Social sensing from street-level imagery: a case study in learning spatio-temporal urban mobility patterns. *ISPRS J. Photogramm. Remote Sens.* 153, 48–58. <https://doi.org/10.1016/j.isprsjprs.2019.04.017>.