Transactions in GIS

WILEY

Transactions in GIS

RESEARCH ARTICLE OPEN ACCESS

Social Inequality in Sight: Exploring Urban Visual Perception and Sentiment Across Income Levels

Ketong Shen | Jian Liu | Xintao Liu

Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China

Correspondence: Xintao Liu (xintao.liu@polyu.edu.hk)

Received: 25 September 2025 | **Revised:** 13 February 2026 | **Accepted:** 24 February 2026

Keywords: computer vision | interpretable machine learning | residents' sentiment | social inequality | urban visual perception

ABSTRACT

Visual perception and sentimental expression are interconnected cognitive processes that shape environmental interaction, functioning at different levels across populations. Despite their inherent connection, prior studies have examined them in isolation, focusing on physical environments and overlooking diverse populations. To address this gap, we investigate the connection between urban visual perception and residents' sentiment, with particular attention to income disparities and social inequalities. Deep learning models are used to analyze street view images and social media data to quantify visual perception and sentiment, while separate CatBoost models are trained for each income-level group and compared using interpretability methods. Our findings show that only moderate visual perception enhances sentiment, while low or high perceptual stimulation reduces sentiment. Furthermore, although economically vulnerable residents are exposed to lower-quality perceptions, enhancements in street conditions and income yield greater improvements. These results highlight social inequalities in the urban environment and suggest that unplanned urban sprawl may harm overall well-being.

1 | Introduction

According to the United Nations World Happiness Report, life satisfaction and happiness are essential measures of national well-being, reflecting progress and success of a country (Helliwell et al. 2023). At the individual level, prior studies show that residents' sentiments are closely linked to happiness and life satisfaction, with positive emotions enhancing subjective well-being and negative emotions such as sadness and anxiety having detrimental effects (Rui 2023; Ma et al. 2024). These sentiments are shaped through everyday interactions between people and their surrounding environments, making them a meaningful reflection of place-based well-being (Fredrickson 1998). Traditionally, this field emphasizes geographic variations in happiness across countries or regions (Easterlin 1995; Wang et al. 2022). With rapid global urbanization and population growth posing challenges to social equality and sustainable development, researchers are increasingly focusing on the spatial heterogeneity of sentiment and well-being within cities.

Ongoing research has documented various factors influencing individual happiness, including economic status (Easterlin 1995; Easterlin 2001), personal characteristics (Alesina et al. 2004; Dolan et al. 2008), and the built environment (Tella et al. 2003; Pfeiffer and Cloutier 2016). Among these, the street-level built environment plays a crucial role in shaping residents' sentiments in urban life. For instance, urban greenery helps restore mental fatigue and alleviate stress, thereby reducing negative emotions and benefiting mental health (Grahn and Stigsdotter 2010; Beute et al. 2023).

However, prior work typically adopts a mechanical assessment of street-level environments, relying on semantic segmentation of Street View Images (SVIs) to quantify physical elements (Rui 2023; Liu et al. 2025). Although valuable, they ignore the inherent connection between sentimental expression and residents' subjective experiences toward the urban environment, known as urban visual perceptions. These perceptions capture cognitive responses that are strongly linked to emotional behaviors, as evidenced by findings in neuroscience and psychology (Dolan 2002; Pessoa 2008). Acknowledging this connection is essential for developing a more comprehensive understanding of residents' sentiment. Moreover, vulnerable residents often endure poorer living conditions and degraded street environments, which may influence not only their emotions but also how they visually perceive and interpret their surroundings (Shaw 2004). Their experiences may be disproportionately affected by the street environment, requiring a more detailed and focused analysis to capture these disparities (Mitchell et al. 2015; Lucas 2012). This suggests the necessity of considering social inequality in the analysis of the perception–sentiment relationship.

In response to these research gaps, this study investigates the relationship between urban visual perception and residents' sentiment, with particular attention to the economically disadvantaged residents. Three specific research questions (RQs) are proposed and answered in this study: (1) Do residents of different income levels experience varying qualities of urban visual perceptions? (2) How does urban visual perception influence sentiment, and do its effects differ between low-income and high-income groups? (3) What role does income play in the relationship between perceptions and sentiment?

To answer these questions, an analytical framework is developed to integrate social media data with machine learning-based assessments of street-level environments. Specifically, geotagged Twitter (X) data are adopted to assess residents' sentiments, while urban visual perceptions are measured using SVIs. To better capture the experiences of local residents, major housing estates (MHEs) serve as the study units and are categorized by median monthly household income (MMHI) to facilitate comparative analysis. Finally, separate nonlinear regression models are developed for each group and explained relying on interpretable techniques. This framework enables us to examine and compare the contributions of various features, providing insights into how urban visual perceptions contribute to sentiment and how these relationships differ by income level. Building on these insights, this study aims to propose actionable recommendations for more efficiently improving the living environment, enhancing vulnerable people's well-being, and promoting social fairness.

2 | Literature Review

2.1 | The Relationship Between Residents' Sentiment and Urban Perception

As urban populations grow and urban sprawl increases, urban issues such as overcrowding and environmental degradation increasingly diminish the quality of the urban environment, negatively affecting residents' health and well-being. Consequently, researchers have turned their attention to understanding how the built environment influences residents' happiness and improving their experiences of urban spaces. For example, studies evaluating accessibility to open and green spaces show that these areas help restore attention and reduce stress, thereby improving residents' sentiment (Grahn and Stigsdotter 2010). In contrast, factors like high building density (Guo et al. 2022; Li et al. 2023) and inadequate public facilities (Lucas 2012) are associated with increased mental stress and reduced life satisfaction.

However, these studies treat urban environment indicators mechanically, focusing on physical features like point of interest (POI) density or pixel counts in SVIs, while ignoring the subjective human experience of the urban environment. Numerous studies demonstrate the strong relationship between subjective perceptions and human behaviors, including crime patterns (Salesses et al. 2013; Zhang et al. 2021) and urban vitality (Ma 2023; Wu et al. 2023). Thus, given that sentiment is a complex aspect of human behavior, it is likely closely linked to urban perception, which reflects the perceptive experience of the urban environment and captures qualities that traditional physical indicators cannot.

From another perspective, research in neuroscience and psychology also shows that perception and cognition are not purely sensory but deeply entangled with sentiment (R J. Dolan 2002; Ochsner and Phelps 2007). Cognitive–emotional interactions are regulated through highly connected brain regions, called hubs, which facilitate the integration and flow of information across different brain systems (Pessoa 2008). These findings highlight the close connection between sentiment and human perception, motivating our investigation into how urban perceptions influence residents' sentiment.

2.2 | Social Inequality in Urban Environment and Residents' Sentiment

With urbanization, the spatial distribution of the urban environment becomes intertwined with socio-economic inequalities among residents. These inequalities contribute to increased class conflict and reduce the life chances and happiness of vulnerable populations, making them a crucial subject of academic investigation.

In particular, economic disadvantages can negatively impact happiness through disparities in both the physical built environment and perceived neighborhood characteristics (Drukker and van Os 2003; Mouratidis 2020). Compared to neighborhood characteristics, the built environment, such as green spaces, local amenities, and transport accessibility, tends to differ more significantly between poor and affluent communities (Mitchell et al. 2015; Mohai et al. 2009). Such disparities indicate that vulnerable residents are more likely to experience poorer urban environments, which can undermine their well-being and further deepen social inequality.

Despite these challenges, the "equigenesis" theory suggests that improvements in public infrastructure and urban facilities can yield greater benefits for vulnerable groups, as they tend to rely more heavily on public services and open spaces than affluent populations. Urban green space is a prominent example, with extensive evidence showing stronger mental health benefits among individuals with lower socio-economic status (Mitchell et al. 2015; Smith et al. 2016). However, existing studies largely attribute social inequalities in well-being to disparities in physical facilities, while paying limited attention to the role of subjective visual perceptions in shaping unequal urban experiences.

As a fundamental human response to the built environment, visual perception provides a direct means of understanding how vulnerable residents experience and interpret urban space (Salesses et al. 2013; Ito et al. 2024). Accordingly, this study investigates social inequalities embedded in the relationship between urban visual perception and sentiment, with a particular focus on vulnerable populations. We aim to provide guidance for efficiently improving the street environment and enhancing the happiness of vulnerable populations.

2.3 | Sentiment in Social Media

Traditionally, the investigation of happiness and life satisfaction has relied on manual data collection methods, such as interviews and questionnaires (Kahneman and Deaton 2010; Mouratidis 2020). While these methods often yield accurate results, they are labor-intensive and unsuitable for large-scale and spatiotemporal analysis. Nowadays, with the rise of platforms like Twitter and Facebook, social media data has emerged as a valuable resource as it offers easy access and provides rich information on spatial interactions and place semantics (Liu et al. 2015; Shi et al. 2022). Numerous studies demonstrate its utility as crowd-sourced data reflecting the real world and human behaviors, such as predicting pandemics (Chunara et al. 2012; Sinnenberg et al. 2017) and capturing urban activity (Chang et al. 2022; He et al. 2023). Accordingly, residents' sentiment, as an essential component of human behavior interacting with the urban environment, is frequently analyzed using social media data (Zheng et al. 2019; Wang et al. 2022).

Recent advances in Natural Language Processing (NLP) and Deep Learning (DL) have further strengthened the use of social media data for sentiment analysis. DL-based approaches transform textual content into numerical representations and employ predictive models to estimate sentiment probability scores. Due to their ability to achieve high accuracy and offer more reliable outcomes when trained on large datasets, model-based methods have become a common approach for estimating happiness and life satisfaction (Zheng et al. 2019). Therefore, this study employs Twitter data and DL methods to estimate residents' sentiment levels in Hong Kong communities.

3 | Data and Methodology

3.1 | Study Area

The study area for this research is Hong Kong, a densely populated metropolitan city with a unique cultural blend. As of 2023, Hong Kong's population of 7.5 million is concentrated within a $1,100\text{km}^2$ territory. As a cultural melting pot of Asian and Western influences, Hong Kong offers a diverse streetscape and varied lifestyle practices, making it an ideal case for studying the link between urban visual perception and sentiment.

Our analysis focuses on 540 MHEs, illustrated in Figure 1. These MHEs are defined as residential communities with at least 3000 residents or 1000 domestic households (C&S Department 2021). In the Hong Kong context, MHEs primarily consist of high-density residential estates, including both public rental housing and large-scale private housing developments. They function as relatively self-contained residential units with shared facilities and relatively homogeneous housing conditions, making them an appropriate spatial unit for examining socio-economic differences in everyday urban experiences. Accordingly, MMHI at the MHE level, as reported in government census data, is used as a reliable indicator to classify MHEs into income groups. Approximately 65% of Hong Kong citizens currently reside in MHEs, underscoring their representativeness of local residential patterns and their widespread use in urban studies (Sit et al. 2025; Lou et al. 2024). By focusing on MHEs, we minimize the irrelevant data, such as tweets from tourists at popular travel sites, allowing us to concentrate on residents' sentiment.

To analyze how the urban environment surrounding residential areas affects sentiment, we create different buffer zones around MHEs to evaluate street environments and urban visual perceptions at varying scales. Smaller buffers are inadequate for capturing residents' potential activity spaces, while larger buffers risk diluting the unique streetscape features within each MHE (Holliday et al. 2017). Thus, based on the typical scope of resident activities, we select buffer distances of 500, 800, and 1000m to identify the most appropriate neighborhood scale.

3.2 | Data Collection

In this study, we construct a cross-sectional dataset integrating Twitter data, SVI data, POI data, and sociodemographic characteristics information. The dataset contains 20 common factors summarized in Appendix Table A1, including their data sources, descriptions, and references, along with six urban visual perception factors. We summarize the values of these factors at the MHE scale to represent the overall characteristics of each MHE community.

3.2.1 | Twitter Check-In Data

We utilize geotagged Twitter data from Hong Kong as a primary source of sentiment information. To build the Twitter database, we employ Twitter's streaming API to collect all tweets containing precise latitude and longitude information within the boundaries of MHEs. No topical keywords were used in the data collection process, allowing the dataset to reflect residents' spontaneous and general sentiment without topic-specific bias. The dataset includes information such as timestamp, tweet text, user ID, language, and geotags.

To ensure data quality, we perform extensive data cleaning following a standard filtering process, including removing repetitive advertisement messages, as well as deleting hyperlinks and special characters (e.g., #, @xxx) (Hao et al. 2023). After processing, our dataset includes 61,080 tweets posted between January 1, 2020, and October 30, 2022. Approximately 70% of the tweets are in English or Chinese, and their number and distribution align with findings from previous studies on Twitter data in Hong Kong (Hao et al. 2023; Chang et al. 2022). Due to privacy protections and data access constraints imposed by the Twitter (X) API, the available dataset is restricted to anonymized user identifiers and tweet-level metadata. Detailed individual-level

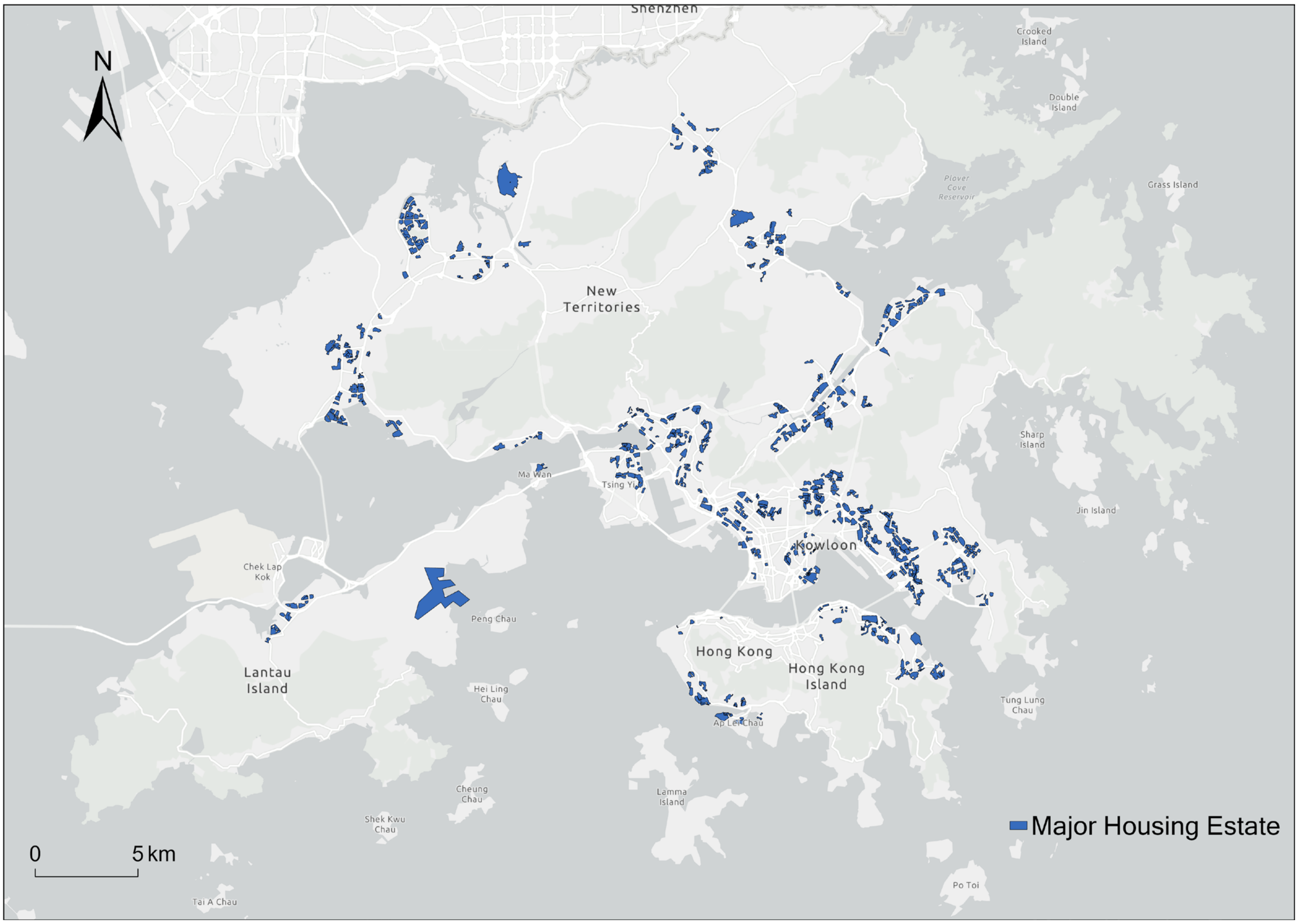

FIGURE 1 | Spatial distribution of MHEs in Hong Kong.

demographic attributes, such as age, gender, or income, are not accessible for geolocated tweets.

3.2.2 | Street View Images

We collect 414,436 SVIs from Google Platform between 2020 and 2022, sampled at 50m intervals along Hong Kong's street network. Figure 2 illustrates the spatial distribution of the sample points across the city. Due to Hong Kong's land-use patterns, where over 60% of the area is designated as country parks, the sample points exhibit spatial clustering, justifying our focus on MHEs for this analysis. At each sample point, we capture four SVIs with a resolution of 640 × 640 pixels, facing directions of 0°, 90°, 180°, and 270°, to provide a comprehensive view of the surrounding street environment.

3.2.3 | Census and POI Information

Population and income data, as sociodemographic characteristics, are drawn from the 2021 Census (C&S Department 2021) to support the analysis of residents' sentiments and tweeting behaviors (Winters and Li 2017; Duan et al. 2022). We also collect POI data, detailing geographic location, address, and type, from Amap (https://lbs.amap.com/api) for analysis. In total, our dataset includes 299,692 POIs, categorized into 21 main classes. For this research, we focus on POI classes relevant to public services and daily convenience, such as public facilities, commercial services, leisure services, education facilities, and financial services. These POI classes are frequently used in studies on sentiment and have been shown to strongly correlate with happiness (Benita et al. 2019; Dong et al. 2024). Finally, we calculate the POI density within MHEs using different buffer zones for analysis.

3.3 | Methodology

We develop a comprehensive framework to analyze the relationship between residents' sentiment and urban visual perceptions, as outlined in Figure 3. In this framework, residents' sentiment is selected as the dependent variable, while independent variables include sociodemographic characteristics, built environment indicators, objective street environment features, and subjective urban visual perceptions. Following prior studies, residents' sentiment is widely used as a behavioral and perceptual proxy for subjective well-being and happiness, which are treated interchangeably. Residnts' sentiment is derived from sentiment analysis of Twitter data, street environment features are quantified from semantic segmentation of SVIs, and urban visual perceptions are predicted using a prediction model trained on the Place Pulse 2.0 dataset. To assess the differences between economically vulnerable and affluent residents, MHE

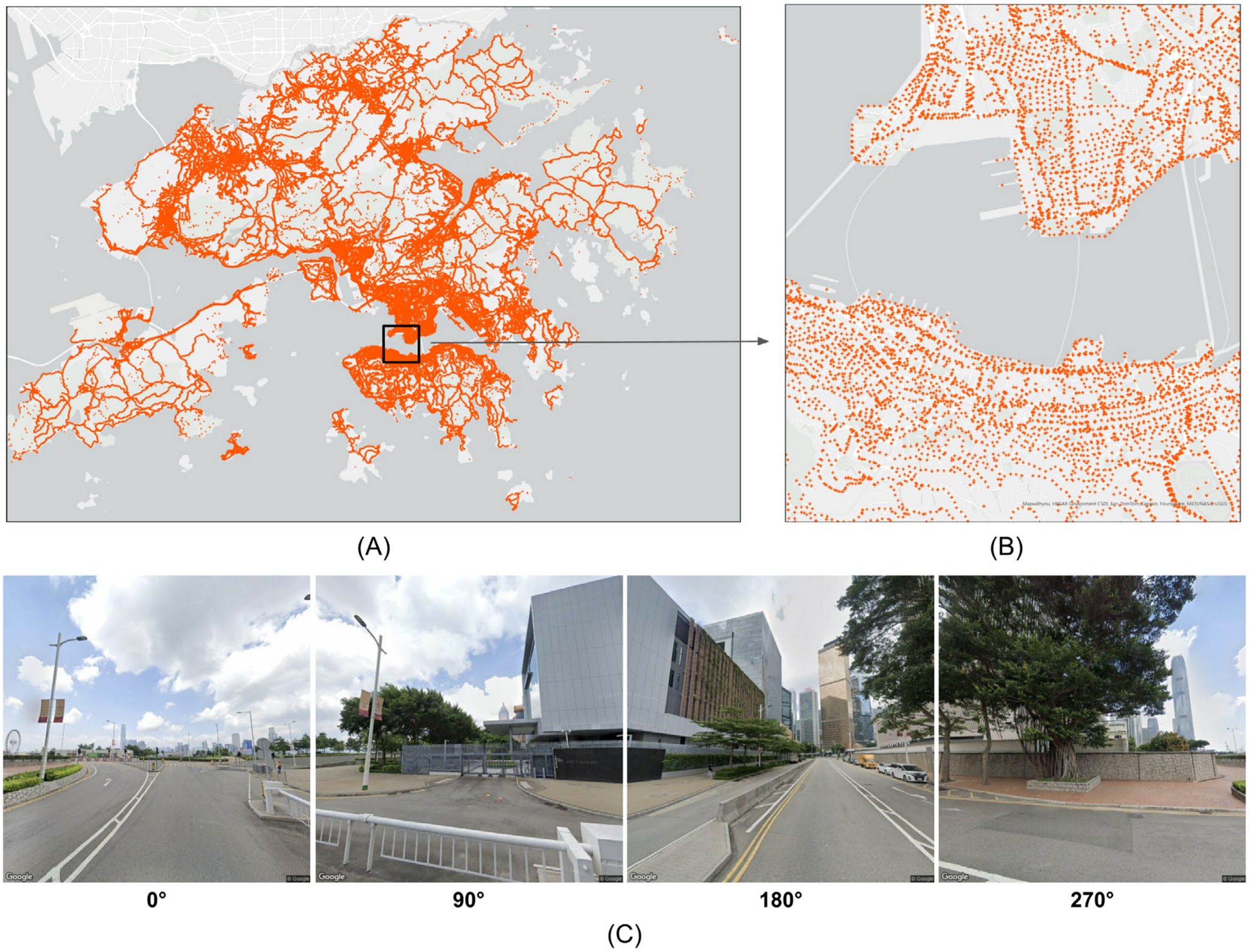

FIGURE 2 | Process for downloading SVIs: (A) Sample points distributed along the street network in Hong Kong, (B) Street sampling points spaced at 50 m intervals, (C) Example SVI captured in four directions at a sample point.

communities are categorized into two groups based on income levels. CatBoost is employed to separately train and test models for the two groups. Finally, we use SHapley Additive exPlanation (SHAP) and Partial Dependence Plots (PDP) for interpretable analysis, identifying the contribution of variables to the predicted sentiment scores.

3.3.1 | Tweet Sentiment Analysis

To calculate tweet sentiment scores, we use TweetNLP for sentiment analysis (Camacho-collados et al. 2022). Trained specifically on Twitter data, TweetNLP is well-suited for predicting sentiment in our multilingual dataset. Its design accounts for multilingual Twitter text, mitigating biases that might otherwise arise from linguistic and cultural variations. By learning general-purpose multilingual representations, it achieves high accuracy across multiple languages (Hu et al. 2020). It classifies sentiment into three categories: positive, neutral, and negative. Sentiment scores are computed by subtracting the probability of a negative sentiment from the probability of a positive sentiment, resulting in a scale from −1 (most negative) to +1 (most positive). These individual tweet scores are then aggregated by averaging the sentiment scores of all geotagged tweets within each MHE, providing a community-level sentiment measure.

3.3.2 | Objective Street Environment Measurement

Measuring the street environment relies on the Pyramid Scene Parsing Network (PSPNet), a deep convolutional neural network trained on the ADE20K dataset, to segment SVIs (Zhao et al. 2017; Zhou et al. 2017). This segmentation enables us to quantify the street environment by calculating pixel ratios for key environmental features, such as sky, buildings, and trees. Since the ADE20K dataset includes 150 categories spanning both indoor and outdoor elements, we follow established methodologies and consolidate segmentation outputs into indicators of street visual quality, such as enclosure, walkability, vitality, and street obstacle (Rui 2023; Ma et al. 2024; He et al. 2023). To further capture the richness of visual information in the street environment, we calculate Shannon Entropy of the top 50 segmentation categories to reflect the variety of visual elements residents encounter in their surroundings (Liu et al. 2022; Chen et al. 2024). The results are then summarized at the community level by mapping them to their corresponding MHEs with different buffer zones.

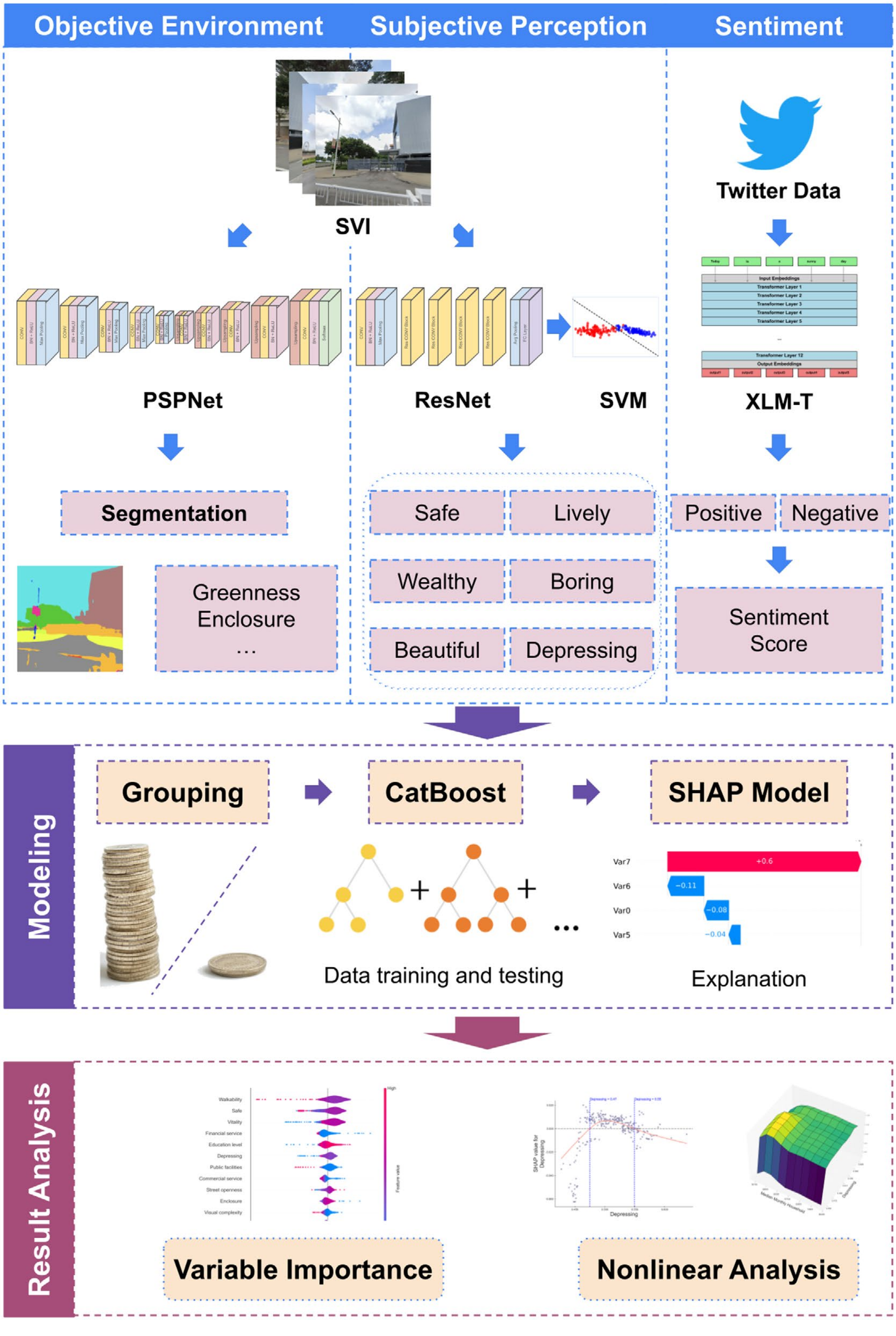

FIGURE 3 | Methodology framework of the study. The sentiment score serves as the dependent variable.

3.3.3 | Subjective Visual Perception Measurement

In this study, we develop our newly trained urban visual perception prediction models for six perception categories using supervised classification. The models are trained using the public Place Pulse 2.0 dataset as labeled training data, covering the dimensions of safe, lively, wealthy, beautiful, boring, and depressing (Dubey et al. 2016).

3.3.3.1 | Training Data Processing. Following the methodology proposed by Salesses et al. (2013), we first convert pairwise comparison outcomes in Place Pulse 2.0 into perception

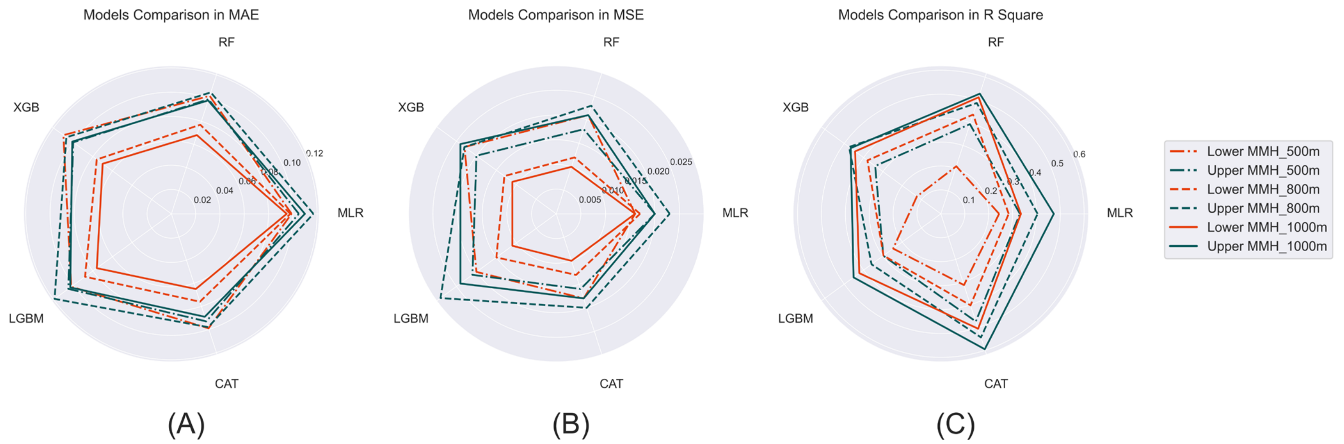

FIGURE 4 | Comparison of evaluation results for different machine learning models. Panels (A), (B), and (C) display the results for MAE, MSE, and (pseudo) R^2, respectively. Across all plots, the model utilizing MHEs with a 1000m buffer (solid line) and the CatBoost algorithm demonstrates the best performance, achieving the lowest MAE and MSE while attaining the highest pseudo-R^2.

scores on a 0–10 scale. Then, due to the inherent subjectivity and uncertainty in human perception, we formulate the prediction task as a binary classification problem instead of a regression task, which ensures more stable categorization of perceptions within intervals (Ordonez and Berg 2014).

For each perception category v, we compute the mean μ_v and standard deviation σ_v of perception scores. Samples with scores higher than $\mu_v + \sigma_v$ are labeled as positive, while samples with scores lower than $\mu_v - \sigma_v$ are labeled as negative:

$$y_i^v = \begin{cases} -1 & \text{if } Q_i^v < \mu_v - \sigma_v \\ 1 & \text{if } Q_i^v > \mu_v + \sigma_v \end{cases} \quad (1)$$

here, positive labels (1) represent strong perceptions (e.g., very beautiful), while negative labels (−1) indicate weak perceptions for the corresponding category.

3.3.3.2 | Prediction Model Construction. In the training process, visual features are extracted from SVIs in the Place Pulse 2.0 dataset using a pretrained ResNet-50 network, which converts each 2D image into a 1D feature vector (He et al. 2016). We use these feature vectors, together with the corresponding perception labels, as inputs to train a separate Support Vector Machine (SVM) classifier for each perception category v. Then, the trained classifiers are applied to the study-area SVI dataset. In this process, SVIs are similarly converted into feature vectors and classified to produce the probability that a given image belongs to the positive perception class (e.g., very beautiful or very safe). This probability is used as the predicted perception score (Zhang et al. 2018).

The training data and prediction framework adopted in this study follow widely used approaches in urban visual perception research. To evaluate model performance within the Hong Kong urban context, we restrict the validation samples to Place Pulse 2.0 SVIs geotagged in Hong Kong, comprising 15,641 human pairwise comparisons. Model performance is evaluated using pairwise agreement accuracy. This metric examines whether the model correctly reproduces human pairwise judgments. Using this metric, the accuracies range from 66.80% to 73.22% across the six perception categories. Given the inherently subjective nature of visual perception and the difficulty of the pairwise prediction task, these results indicate robust and comparable performance relative to existing perception prediction models (Dubey et al. 2016; Yu et al. 2026).

Finally, perception scores are aggregated at the community level by mapping each SVI to its MHE buffer and averaging scores across categories. Appendix Figure C1 illustrates the spatial distribution of six urban visual perceptions, summarized within MHEs using a 1000m buffer.

3.3.4 | Regression Model for Analysis

We apply CatBoost, an optimized gradient boosting algorithm that builds ensemble predictors, to analyze the relationship between urban visual perception and residents' sentiment (Prokhorenkova et al. 2018). CatBoost is selected through an empirical comparison with several commonly used regression methods and is chosen because it achieves the best predictive performance on our data. This data-driven model selection strategy is commonly adopted in related urban analytics studies (Lian et al. 2026; Chen et al. 2025). Specifically, we evaluate CatBoost against multiple linear regression (MLR), random forest regression (RF) (Breiman 2001), eXtreme Gradient Boosting (XGBoost) (Chen and Guestrin 2016), and Light Gradient-Boosting Machine (LGBM) (Ke et al. 2017). All models are trained on 70% of the data, with the remaining 30% reserved for testing, and optimized hyperparameters are applied to ensure a fair comparison.

In Figure 4 and Appendix Table B2, we present the results of model training. The "Total" group represents the outcomes when the entire dataset is used in the training process. In the empirical comparison, model performance is assessed using Mean Squared Error (MSE), Mean Absolute Error (MAE), and (pseudo) R^2 on the testing set. MAE and MSE jointly reflect the magnitude and stability of prediction errors, whereas R^2 captures the model's explanatory power. Among the evaluated models, CatBoost demonstrates the best performance, achieving the lowest MAE and MSE and the highest pseudo-R^2 across all three test groups. The consistent superiority of CatBoost across metrics indicates that it provides both accurate and robust regression, making it well-suited for capturing the non-linear relationships examined in this study. Furthermore, we observe that larger buffer zones for summarizing street environments and perceptions improve model performance, with the 1000m buffer zone producing the most favorable results. Based on these findings, we select the CatBoost model with a 1000m buffer zone for further analyses and experiments. Because all variables are standardized and the gradient-boosted methods inherently perform recursive feature selection, our CatBoost models effectively mitigate multicollinearity issues (Chen and Guestrin 2016).

3.3.5 | Analysis of Nonlinear Relationship

Machine learning methods often function as "black boxes," making their predictions difficult to interpret. To enhance the interpretability of our model, SHAP is primarily used to quantify feature importance and explain model predictions at both global and local levels, demonstrating the contribution of each feature to model predictions. A positive SHAP value indicates that a feature increases the predicted sentiment score, whereas a negative value indicates a suppressing effect (Lundberg and Lee 2017). In addition, PDP is employed to visualize the marginal effects of features and their nonlinear relationships with the predicted outcome.

Firstly, we conduct both global and local SHAP analyses to elucidate the prediction process and outcomes. Global SHAP analysis is performed by averaging the absolute SHAP values across all observations to assess the overall importance of each variable in the prediction model. To facilitate comparison across variables, these aggregated values are further normalized as percentage shares of the total contribution. Local SHAP analysis is used to examine how the contribution of individual urban visual perception variables varies with their feature values, allowing us to explore detailed and nonlinear relationships between the street environment and residents' sentiment.

Based on Shapley values in Game Theory, SHAP calculates the contribution ϕ_i of each feature i as follows:

$$\phi_i = \sum_{S \in F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \left[f_{S\cup\{i\}}\left(x_{S\cup\{i\}}\right) - f_S(x_S) \right] \tag{2}$$

here, F is the set of all features, $f_{S\cup i}$ represents the model trained with feature i included, and f_S represents the model trained without it. The terms x_S and $x_{S\cup i}$ denote the input feature values for sets S and $S \cup i$, respectively.

Then, PDP is used to analyze the marginal effects of urban visual perceptions (Greenwell et al. 2018). The partial dependence function $\hat{f}_P$ is calculated across the dataset using a Monte Carlo approach:

$$\hat{f}_P(x_P) = \frac{1}{n} \sum_{i=1}^{n} \hat{f}\left(x_P, X_C^{(i)}\right) \tag{3}$$

In this equation, $\left\{X_C^{(i)}\right\}_{i=1}^{n}$ represents the observed values of features not of interest, while x_P contains the features of interest. We apply two-way PDP ($|P| = 2$), pairing income level with urban visual perceptions to analyze their interaction marginal effects on the prediction of residents' sentiment.

4 | Experiments and Results

In subsequent experiments, we summarize the objective street environment and urban visual perceptions within MHEs using a 1000m buffer zone and then apply CatBoost to develop nonlinear regression models.

4.1 | Spatial Patterns of Residents' Sentiment

Figure 5 presents the spatial distribution of sentiment scores. The dataset's average sentiment score is 0.311, reflecting an overall positive sentiment. In terms of spatial distribution, inland residential communities, such as the central area of Kowloon (Figure 5C), tend to exhibit lower sentiment levels. Conversely, communities in traditionally affluent areas, such as Hung Hom in eastern Kowloon and the Eastern District (situated in the lower-left corner of Figure 5D), exhibit higher sentiment levels.

We first conduct a global Moran's I analysis to measure the spatial clustering of residents' sentiment. The result shows Moran's I value of 0.5420 with $p < 0.001$, demonstrating significant spatial clustering of sentiment. To pinpoint specific clusters, we then perform Local Moran's I (LISA) analysis (Anselin 1995). As shown in Figure 6, "low–low" clusters are primarily located in underdeveloped areas, such as Sham Shui Po and Kwun Tong. In contrast, "high-high" clusters are concentrated in affluent coastal regions, such as Sha Tin and the Eastern District, where property prices are higher and poverty rates are below the Hong Kong average (C&S Department 2024). The findings from the LISA analysis highlight a potential spatial alignment between sentiment clustering and income levels, with low-sentiment clusters in economically disadvantaged areas and high-sentiment clusters in wealthier regions.

4.2 | Comparison of Different Income-Based MHE Groups

By examining spatial patterns, we observe potential differences in residents' sentiment between lower-income and higher-income groups. However, these patterns provide only qualitative insights. To enable quantitative analysis, we divide the MHE dataset into two groups based on MMHI: MHEs above the overall median (upper MMHI group) and those below the median (lower MMHI group). The median MMHI is 29,845HK$, closely matching Hong Kong's official MMHI of 30,000HK$ (C&S Department 2024), which supports the representativeness of our MHE data. Figure 7 presents the spatial distribution of the two groups, with the lower and upper MMHI groups shown in brown and dark green, respectively. The accompanying histogram shows that the upper MMHI

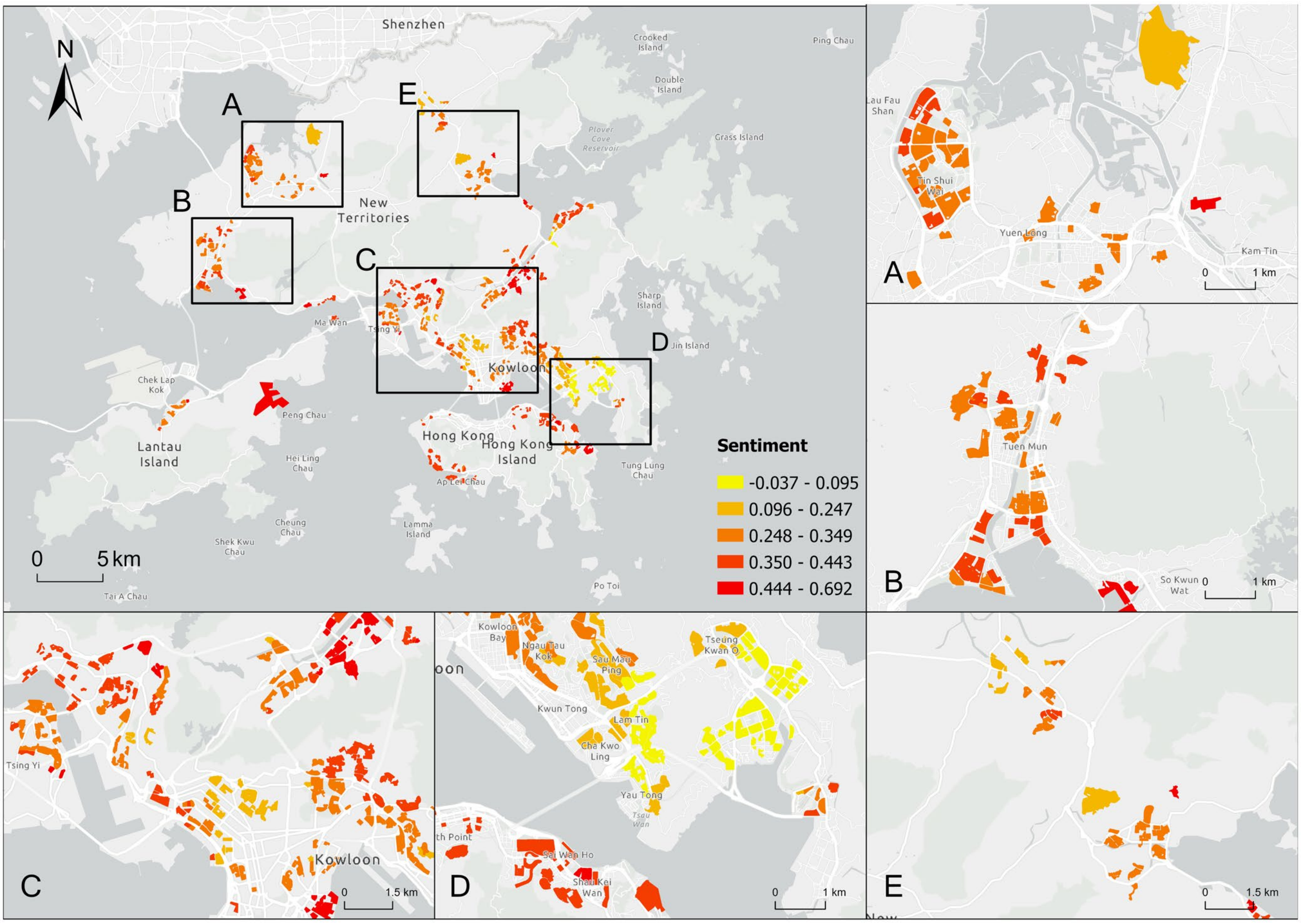

FIGURE 5 | Spatial distribution of summarized sentiment scores into MHEs with 1000m buffer. (A–E) show the details of main residential areas in Hong Kong: (A) Tin Shui Wai and Yuen Long, (B) Tuen Mun, (C) Kowloon Area (including Sham Shui Po, Kowloon City, Wong Tai Sin, and Yau Tsim Mong), (D) Kwun Tong and Sai Kung (including Tseung Kwan O), (E) Tai Po.

group follows a right-skewed income distribution, as commonly observed in the typical income distribution. Consistent with the LISA results in Figure 6, areas such as Sham Shui Po and Kwun Tong exhibit higher concentrations of lower-MMHI MHEs, corresponding to "low–low" sentiment clusters.

Next, we compare perceptions and sentiment levels between the two income groups to explore whether social inequality is reflected in both the living environment and well-being. Figure 8 shows that the upper MMHI group perceives more beautiful but more boring street environments, likely due to better greenery, uniform architectural style, and lower building density in wealthier areas (Zhang et al. 2018). Furthermore, streetscapes in the upper MMHI group are perceived as safer and less depressing, reflecting the superior street environments often associated with wealthier communities. In contrast, the lower MMHI group experiences a livelier street environment, possibly due to bustling yet chaotic streets lined with small shops in low-income communities. These observations highlight social inequalities, where wealthier residents tend to enjoy a higher-quality urban environment and perceptions.

Interestingly, despite this, the average sentiment score shows no significant difference between the two groups. In fact, the lower MMHI group exhibits a marginally higher average sentiment score—just 0.0002 more than the upper MMHI group. This result contrasts with the patterns observed in the spatial analysis. Thus, we further explore the differences by training separate nonlinear models for the upper and lower MMHI groups.

4.3 | Nonlinear Models for MHE Groups

4.3.1 | Global Contribution Analysis

After analyzing the differences in urban visual perceptions and sentiment levels, our results reveal notable social inequalities between the two income-based groups. However, these disparities are not fully supported by basic comparisons. To investigate, we train separate nonlinear models for the upper and lower MMHI groups and then apply SHAP to evaluate the global contribution of independent variables in each model.

Figure 9A,C displays each feature's percentage of the total global SHAP value, highlighting the relative importance of each feature in the model. Figure 9B,D shows violin plots that visualize the variation in SHAP values for each feature across all samples. For the lower MMHI group, depressing, safe, and beautiful perceptions rank first, fourth, and seventh, with

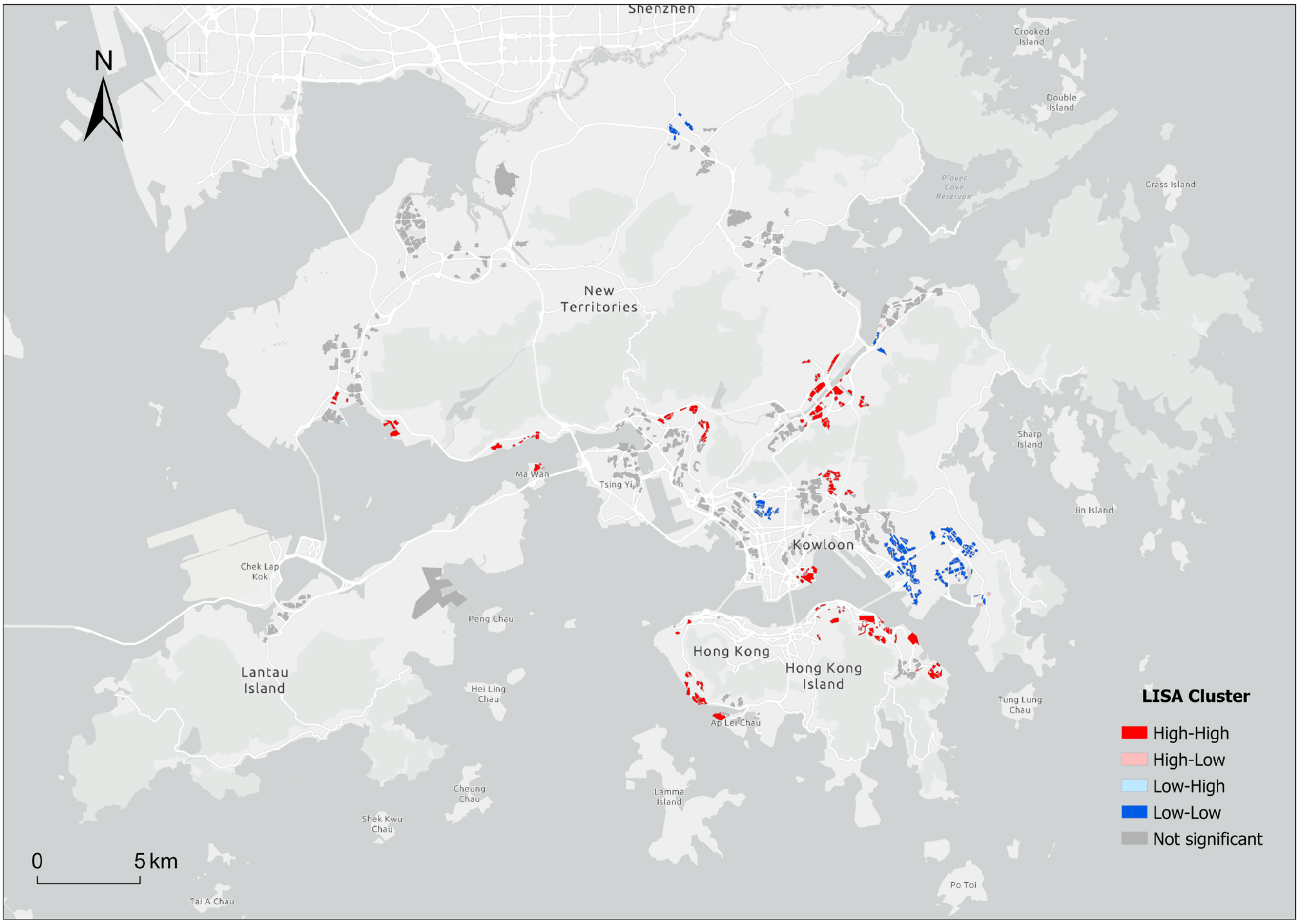

FIGURE 6 | LISA result for sentiment scores aggregated within MHEs using a 1000 m buffer.

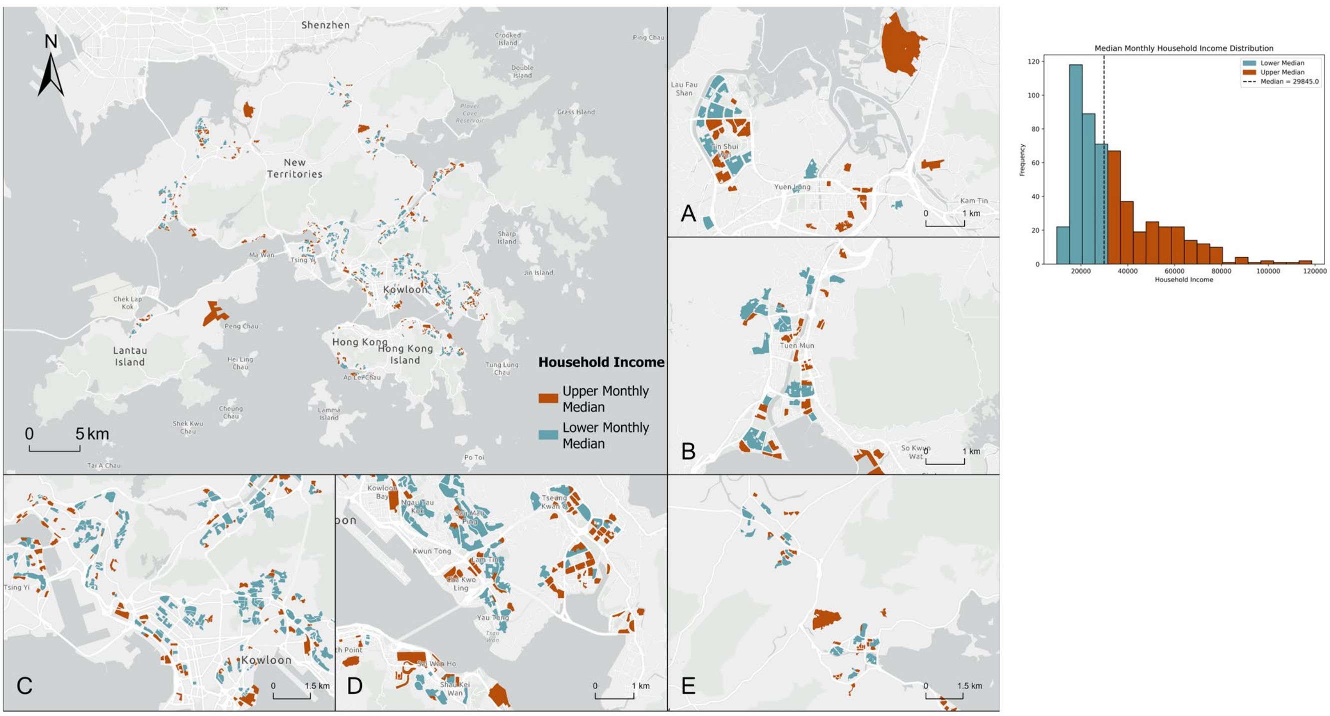

FIGURE 7 | Spatial distribution of MHEs categorized by MMHI levels: The left subfigure highlights MHEs with MMHI below the median in brown, while those above the median are shown in dark green. (A–E) correspond to the main residential areas, consistent with Figure 5. The right subfigure presents the histogram of MMHI distribution across the entire dataset.

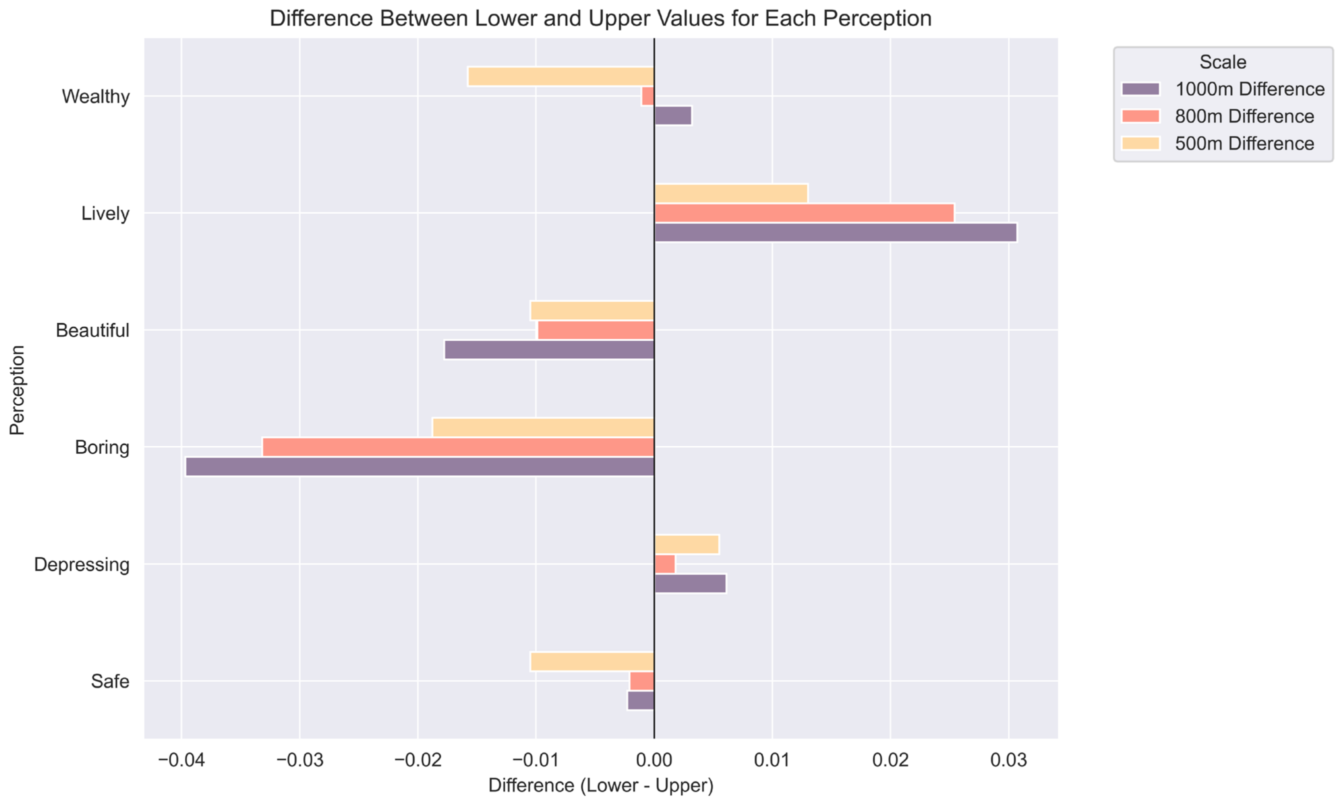

FIGURE 8 | Differences in urban visual perception values between two income-level groups, where a negative value indicates that the upper MMHI group has a higher value than the lower MMHI group. Purple, orange, and yellow colors represent comparisons made in the 1000, 800, and 500m buffer zones, respectively.

percentage SHAP contributions of 8.42%, 6.21%, and 5.78%, respectively. Relative to findings in prior studies, these values are notably large, particularly considering the inclusion of 26 variables in our model (Rui 2023; Jin et al. 2025). The violin plot highlights the negative contribution of depressing perceptions and the pronounced positive contribution of beautiful perceptions, highlighting the critical role of perceptions in predicting sentiment in lower-income communities. Sea view emerges as the second most significant factor (7.60%), showing higher proportions of sea view in the street environment, correlating positively with sentiment levels.

For the upper MMHI group, the factors influencing sentiment differ markedly. Walkability and vitality rank very high, with percentage SHAP contributions of 12.49% and 10.05%, respectively. Low walkability has a minimal impact, with SHAP values clustered around 0.02, but excessive walkability shows a negative contribution to the prediction of sentiment. This outcome may be attributed to suburban small roads, which often feature large pavements that enhance walkability but can also evoke fear of crime, thereby reducing residents' sentiment (Cheng and Smyth 2015; Pfeiffer and Cloutier 2016). Vitality enhances sentiment, and its contribution is more pronounced in the upper MMHI group. This is likely because its positive contribution is diminished in lower-income areas, where dilapidated and overcrowded streets reduce the quality of living space for vulnerable residents (Mouratidis and Yiannakou 2022). Safe perception remains a significant contributor in the upper MMHI model, with a percentage of 10.10%. Conversely, other perception factors play a considerably smaller role, with all contributing under 5%.

To facilitate comparison, we create a bipartite graph (Figure 9E) that contrasts the rankings of factor contributions between two groups, with perception factors represented in black and common factors in gray. Safety consistently ranks high in both models, emphasizing the universal importance of personal security and aversion to crime across income levels. On the other hand, remaining perceptions shift to lower ranks in the upper MMHI group, indicating that sentiment in this group is less associated with perceptions.

4.3.2 | Local Contribution Analysis

Based on the global contribution analysis, urban visual perceptions show a strong contribution to the prediction of sentiment, particularly among lower-income residents. To further explore this nonlinear relationship, we narrow our focus to perceptions in the models. Specifically, we focus on the three highest-ranking perceptions, namely depressing, safe, and beautiful, as features with minimal contributions tend to produce flat trends, making the results less meaningful.

Figure 10 presents the results of the local SHAP analysis. The smoothed trend of SHAP values highlighted by the red curve reveals pronounced nonlinear relationships. The blue vertical markers denote the points at which the contribution shifts,

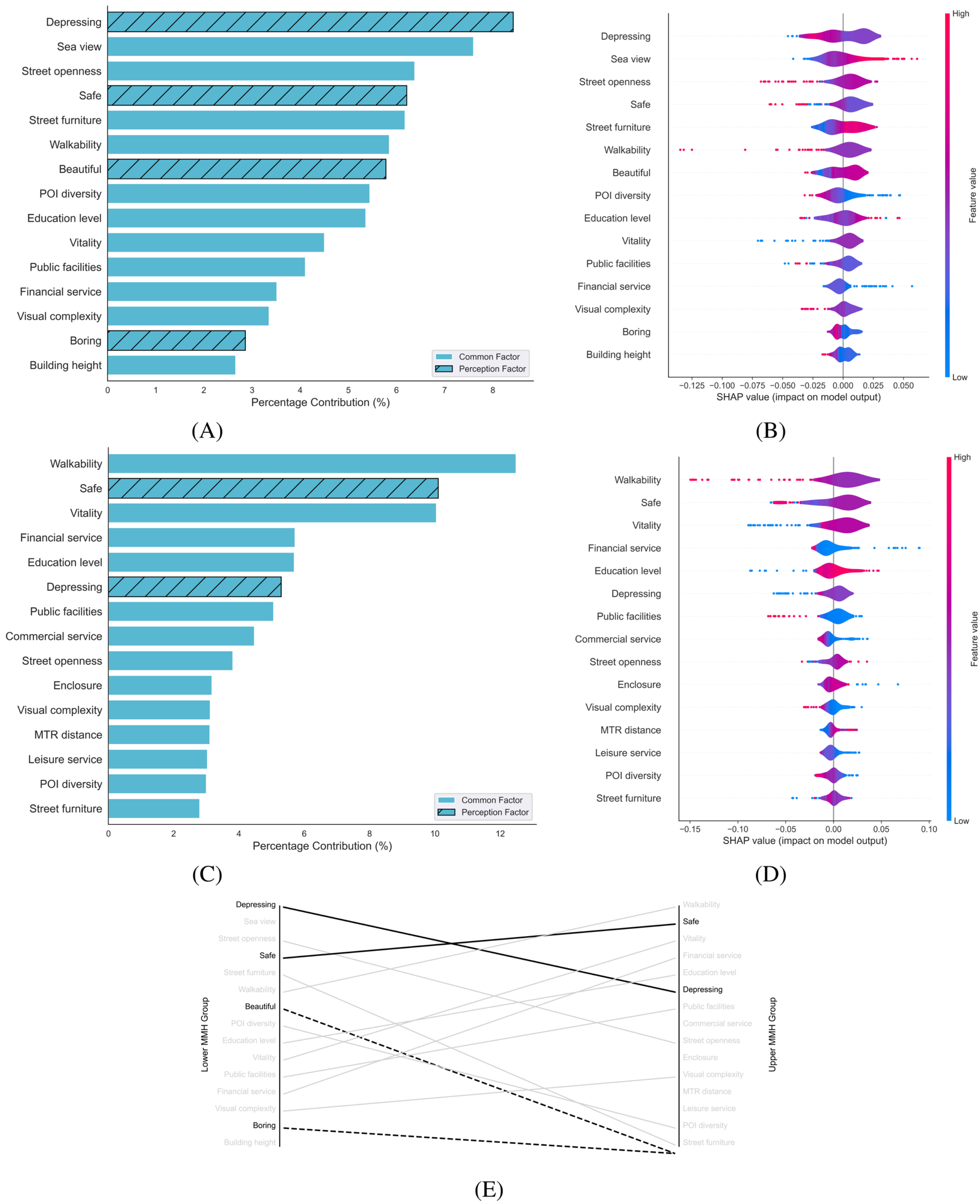

FIGURE 9 | SHAP results for the nonlinear models of the lower and upper MMHI groups: (A) and (B) display the SHAP results for the *lower* MMHI group; (C) and (D) present the results for the *upper* MMHI group. The left subfigures show bar charts representing the percentage SHAP contributions, and the right subfigures display violin plots illustrating the distribution of SHAP values across samples. (E) shows the comparison of rankings of factor contribution between the lower MMHI group and the upper MMHI group.

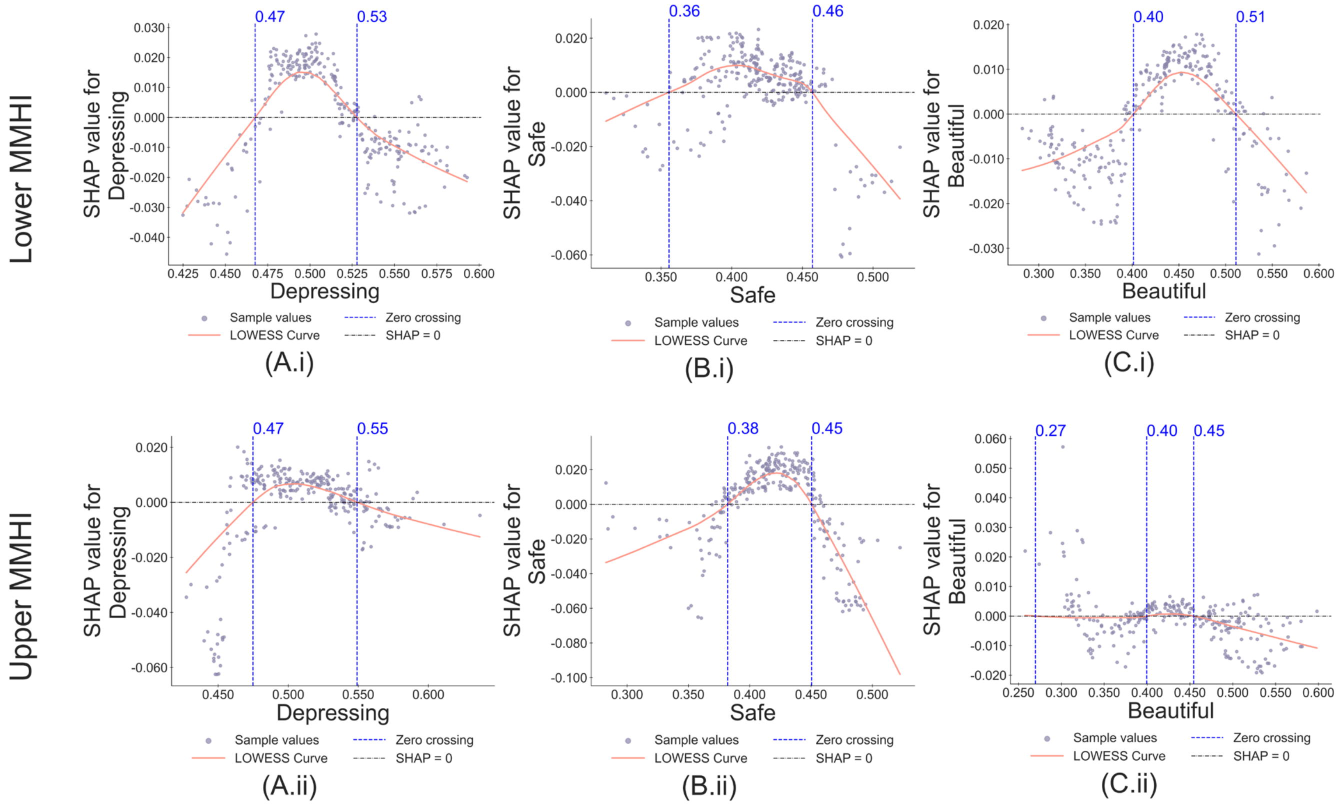

FIGURE 10 | Local SHAP analysis for the three highest-ranking perceptions in the nonlinear models of the lower MMHI group and the upper MMHI group. (A–C) are depressing, safe, and beautiful perceptions, respectively. The red curve represents a LOWESS-smoothed trend of SHAP values as perception scores change. Blue vertical lines indicate the points where the trend crosses zero, marking transitions between positive and negative contributions to the prediction of sentiment.

identifying critical thresholds beyond which the same perception begins to exert an opposite contribution to the prediction of residents' sentiment. Since all variables are standardized during training to align with the requirements of gradient boosting algorithms (Chen and Guestrin 2016; Prokhorenkova et al. 2018), the values of variables in the results are rescaled to their original ranges for interpretation. In the lower MMHI group, perceptions exhibit non-monotonic trends: they show positive contributions at medium levels but negatively contribute to the prediction of sentiment at both low and high levels (Figure 10Ai,Bi,Ci). For example, a depressing visual perception, ranked as the most important factor, is associated with lower sentiment levels when high, as dense buildings and narrow streets obstruct visual openness and sunlight (Guo et al. 2022; Li et al. 2023). Yet, when depressing perception becomes very low, overly empty and spacious streets with few elements can also diminish sentiment. Likewise, safe and beautiful perceptions follow similar patterns, reinforcing the idea that residents tend to prefer a "moderately perceived environment."

For the upper MMHI group, perception-related contributions are smaller. Trends for depressing and beautiful perceptions are less distinct, with most observations clustering near the x-axis ($|\mathrm{SHAP}| < 0.02$). Safety still plays a notable role, suggesting that concerns about security are shared across both groups. Despite the reduced contribution overall, the patterns largely mirror the non-monotonic relationships observed in the lower MMHI group (Figure 10Aii,Bii). This similarity suggests that the preference for a "moderately perceived environment" is consistent across income levels.

4.3.3 | Interaction Analysis of Income and Perceptions

While comparing the two models highlights that urban visual perceptions play a greater role in the sentiment of economically disadvantaged residents, the role of income within the perception–sentiment relationship still needs further analysis. To address this, we employ PDP analysis to assess the interaction marginal effects between income level and perceptions. PDP is based on partial dependence functions, allowing all perception variables to be included in the analysis. The results are shown in Figure 11 for the nonlinear model of the lower MMHI group and Figure 12 for the upper MMHI group.

In the lower MMHI group, as shown in Figure 11A, the contribution of income level exhibits an increasing trend. When MMHI is below approximately 22,000HK\$, income level negatively contributes to the prediction of sentiment, but this negative contribution diminishes as MMHI increases. Once MMHI exceeds 22,000HK\$, the contribution becomes increasingly positive. This trend aligns with prior studies that link recessions and unemployment to reduced well-being and financial success to greater happiness (Dolan et al. 2008; Tella et al. 2003).

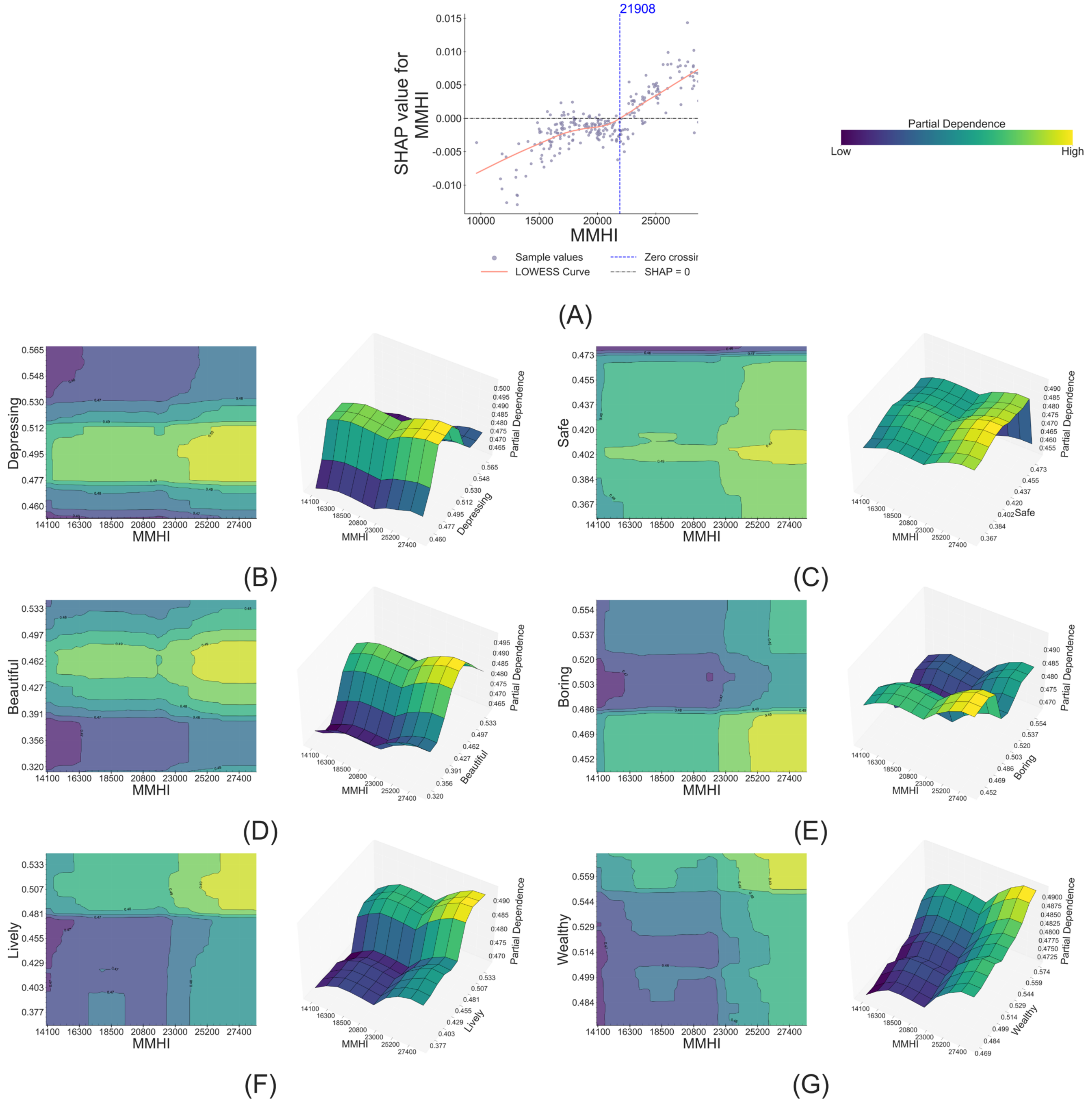

FIGURE 11 | Two-way PDP analysis of income levels and urban visual perceptions within the nonlinear model for the *lower* MMHI group. (A) illustrates the nonlinear contribution of MMHI income to sentiment levels. (B–G) shows the interaction marginal effects between MMHI and various perceptions, including depressing, safe, beautiful, boring, lively, and wealthy, respectively. Yellow areas indicate high partial dependence values, representing positive and strong interaction marginal effects at specific MMHI values and perception scores. Conversely, deep blue areas represent low partial dependence values, highlighting negative or weaker interaction marginal effects.

When urban visual perceptions are incorporated, the trends become more nuanced. Depressing perception exhibits an antagonistic relationship with the income level: as streetscapes become more depressing and income levels remain low, sentiment significantly declines (Figure 11B). A similar antagonistic pattern is observed for boring perception (Figure 11E), suggesting that negative perceptions generally have adverse contributions to the sentiment of residents in the lower MMHI group. On the other hand, positive perceptions, such as beautiful, lively, and wealthy, exhibit synergistic relationships with the income level (Figure 11D,F,G). These results demonstrate that high-quality urban environments, combined with increased income, contribute to enhanced happiness among residents. Safe perception, by contrast, does not display significant interaction with MMHI (Figure 11C), as its consistently positive contribution to sentiment remains unaffected by changes in income. This finding aligns with the results of SHAP analysis, confirming that safety is a universal concern across all income levels.

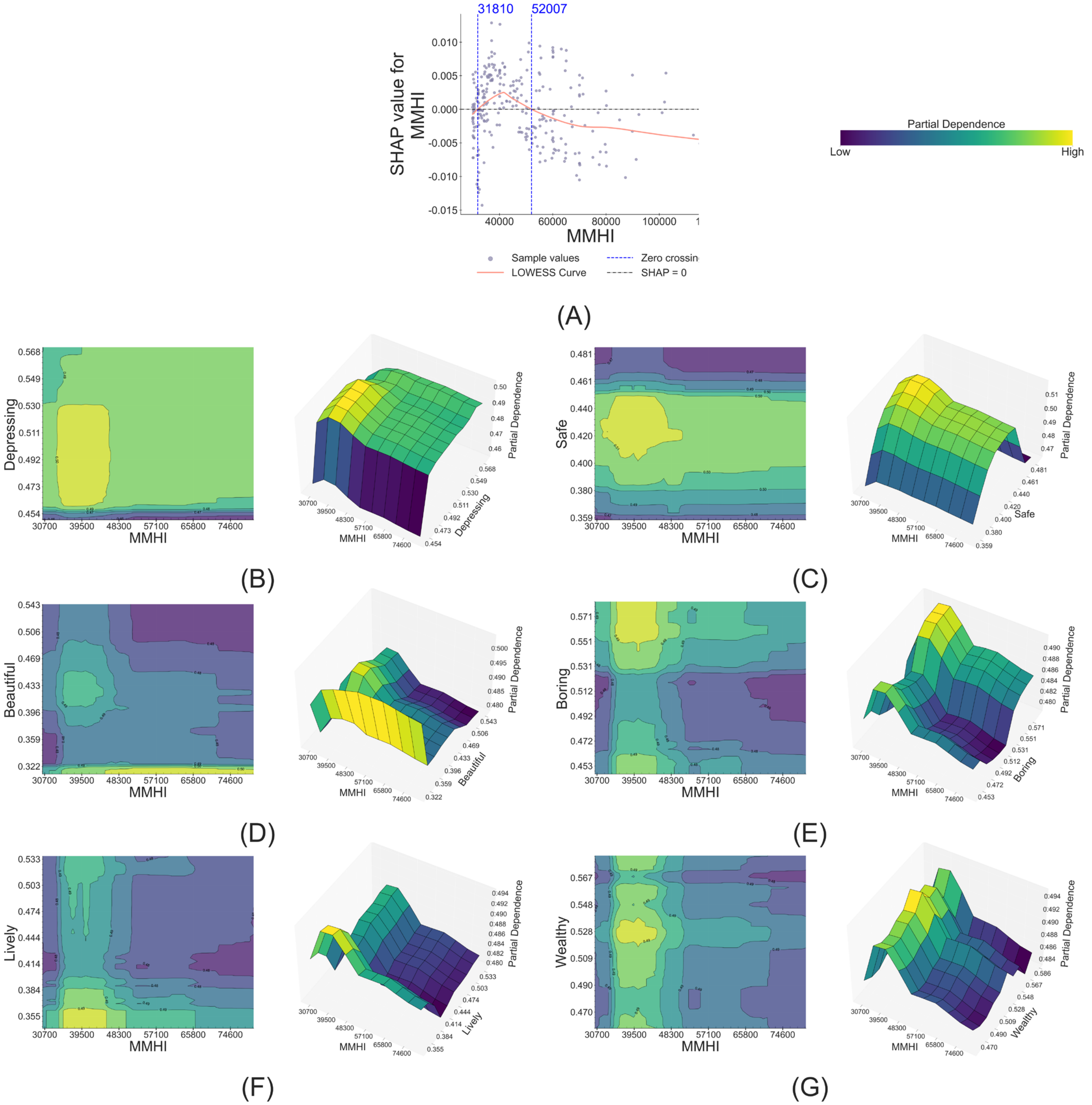

FIGURE 12 | Two-way PDP analysis of income levels and urban visual perceptions within the nonlinear model for the *upper* MMHI group. (A) illustrates the nonlinear contribution of MMHI income to sentiment levels. (B–G) shows the interaction marginal effects between MMHI and various perceptions, including depressing, safe, beautiful, boring, lively, and wealthy, respectively. Yellow areas indicate high partial dependence values, representing positive and strong interaction marginal effects at specific MMHI values and perception scores. Conversely, deep blue areas represent low partial dependence values, highlighting negative or weaker interaction marginal effects.

When MMHI exceeds the median income, the trend shifts significantly. As shown in Figure 12A, the local SHAP values of the income level in the upper MMHI group reveal a decreasing trend. In this affluent group, higher income levels do not enhance happiness; rather, they detract from well-being. This result supports research suggesting that, beyond a certain income threshold, happiness is more closely linked to social relationships than financial success (Diener and Seligman 2004). Interestingly, the interaction marginal effects between income and perceptions in the upper MMHI group are less significant than those in the lower MMHI group. Most perceptions contribute consistently to sentiment regardless of the income level.

5 | Discussion and Conclusion

Visual perception and sentiment are deeply interconnected in human cognition, often integrating to jointly influence behaviors (Dolan 2002; Pessoa 2008). This interplay highlights the complexity of human sentiment and emphasizes the importance of exploring the connection between visual perceptions and sentimental expression in the urban environment. Moreover, considering the uneven urban development that characterizes urbanization, social inequalities highlight the need for greater attention to the well-being of vulnerable communities in the analysis.

In this study, we use Twitter data and SVIs to examine how urban visual perceptions shape residents' sentiment across different income-level groups. To explore RQ1, we categorize communities in Hong Kong into two income-level groups. Our findings reveal that wealthier communities experience significantly better urban visual perceptions, uncovering a notable dimension of social inequality in the urban environment. To gain deeper insights, we train separate nonlinear models for each group, revealing that urban visual perceptions play a more prominent role in sentiment among lower-income residents. Specifically, only urban environments with moderate perception scores, referred to as "moderately perceived environment," positively contribute to the prediction of sentiment, addressing RQ2. When analyzing the role of income levels to investigate RQ3, we find that income growth enhances sentiment among economically vulnerable communities and amplifies the positive contributions of urban visual perceptions. Conversely, for wealthier residents, additional income growth may diminish sentiment levels.

These results further highlight the social inequalities, where vulnerable communities often lack the resources needed to improve their surroundings, even though the quality of the urban environment plays a more crucial role in shaping their well-being. This underscores the importance of prioritizing improvements in street environments and living conditions in vulnerable communities to achieve more equitable well-being outcomes.

5.1 | Urban Visual Perceptions and Residents' Sentiment

In this study, we observe a positive contribution of a "moderately perceived environment" on residents' sentiment. This finding reveals a non-linear relationship between urban visual perception and sentiment: rather than being maximized under extreme conditions, residents' sentiment appears to benefit most from a balanced and contextually appropriate visual environment.

Poor visual perception often reflects deteriorated or disorderly environments, which can evoke stress, insecurity, and negative affect. For instance, perceptions of unsafe environments increase fear of crime and are consistently associated with lower sentiment levels (Cheng and Smyth 2015; Salesses et al. 2013). Similarly, the contribution of beautiful perception increases from negative to positive as the perception score rises, suggesting that residents favor more aesthetically pleasing urban environments, such as urban parks with better greenery and art elements (Grahn and Stigsdotter 2010; Mitchell et al. 2015).

However, excessively high-quality visual perception also does not necessarily enhance sentiment. One possible explanation lies in cognitive overload: environments with very high safety or visual order are often characterized by crowded streets and rich visual stimuli, which may increase mental fatigue and reduce emotional comfort (Mouratidis and Yiannakou 2022). Another explanation relates to social interaction and perceived social inclusion. Highly curated or "over-perfect" streetscapes may emphasize visual order and regulation. Such environments can signal exclusivity or implicit social norms, potentially intensifying feelings of social exclusion or non-belonging, particularly among vulnerable residents. As a result, high visual quality alone may not translate into higher sentiment.

Our results also suggest that unplanned urban sprawl, characterized by single land use, leapfrog development, and low-density patterns, is detrimental to well-being (Calthorpe and Fulton 2001). Both crowded and dilapidated old city centers and disorderly urban expansion areas often feature excessively stimulating perceptions, which diminish residents' sentiments. Thus, our study emphasizes the importance of regulating urban sprawl and designing street environments to foster a "moderately perceived environment." Proper urban planning in the process of urban sprawl can contribute to more harmonious streetscapes, enhancing well-being and achieving sustainable development goals.

5.2 | Social Inequality in Sentiment

Vulnerable residents are more likely to encounter substandard housing conditions, deteriorated living environments, and restricted access to public services (Mouratidis 2020). Our findings confirm this social inequality, as wealthier communities exhibit higher-quality urban environments compared to lower-income communities. This results from the clustering of vulnerable residents in deprived neighborhoods due to limited financial means, a form of self-selection that reflects socio-spatial injustice (Shaw 2004; Sampson 2019).

Despite wealthier residents experiencing superior street environments, our study reveals that urban visual perceptions have a more substantial impact on economically vulnerable residents. This result aligns with the "equigenesis" theory, which suggests that vulnerable groups often gain more from improvements in social, physical, or service environments than affluent groups (Mitchell et al. 2015; Lucas 2012). As an essential component of public services, the street environments, particularly the levels of visual perceptions, also play "equigenesis" roles in the well-being of vulnerable residents, although their contributions are not uniformly positive. To further investigate the role of income level, we examine (1) the association between income and predicted sentiment and (2) the role of income in the perception–sentiment relationship. We find that in the lower MMHI group, income contributes increasingly positively to sentiment as income rises, whereas in the upper MMHI group, additional income gains are associated with diminishing sentiment. This pattern is consistent with the "Easterlin Paradox," which suggests that happiness increases with income at low-income levels but shows weak long-term correlation with income in wealthier contexts (Easterlin 1995; Easterlin 2001). For economically vulnerable residents, limited financial resources are closely linked to emotional distress, and income growth helps satisfy basic needs, thereby enhancing sentiment (Blanchflower and Oswald 2004; Kahneman and Deaton 2010). In contrast, the correlation between income and well-being becomes less in wealthier societies, as illustrated by a reported correlation of 0.45 in the slums of Calcutta compared to 0.13 in the United States (Biswas-Diener and Diener 2001; Diener et al. 1993). It may result from social comparison, which diminishes the happiness gained from additional income (Dolan et al. 2008; Easterlin 2023).

Additionally, our results demonstrate that income growth amplifies the positive association between urban visual perceptions and sentiment for lower-income residents. This further confirms the "equigenesis" role of urban visual perceptions and demonstrates the stronger aspirations of economically vulnerable residents for better living environments. We believe our results highlight the necessity of building equitable street environments and emphasize the importance of prioritizing investments in vulnerable communities.

5.3 | Research Contribution

In this paper, our contributions are threefold. First, our findings demonstrate the crucial role of urban visual perception in shaping residents' sentiments. Specifically, a "moderately perceived environment" positively contributes to higher sentiment, whereas excessive sensory stimulation is associated with lower sentiment. This highlights the potential of unplanned urban sprawl to reduce residents' well-being. Second, rather than treating all residents as a homogeneous group, we account for inherent social inequalities in urban environments. Our results show that urban visual perceptions play a more important role in shaping sentiment among lower-income residents, even though they experience lower-quality environments. This research uncovers social inequality from a perception-based perspective and provides empirical support for the "equigenesis" theory. Third, we investigate the role of income levels in shaping sentiments. For economically disadvantaged residents, income growth not only improves well-being but also strengthens the positive association between urban visual perceptions and sentiment. These results underscore the necessity of prioritizing investment in vulnerable communities to efficiently enhance overall well-being and reduce urban disparities.

Together, these contributions advance the understanding of the non-linear relationship between urban visual perceptions and residents' sentiment, particularly in the context of social inequality. By uncovering the unequal distribution of perceptual experiences and their disproportionate contribution to vulnerable groups, our study emphasizes the need for perception-sensitive and equity-oriented urban design. These insights offer practical guidance for policymakers and planners aiming to improve well-being and promote social fairness through targeted interventions.

5.4 | Policy Implication

The findings of this study have several policy implications. First, our findings underscore the need to control unplanned urban sprawl. In particular, strategies should aim to prevent overdevelopment in city centers and poorly integrated suburban growth by promoting mixed land use, visual coherence, and more balanced streetscapes. Such planning approaches can help create "moderately perceived environments" that support well-being. Second, high-quality urban visual perceptions are particularly important for vulnerable residents. Thus, decision-making and planning departments should focus on improving the living and street environments in lower-income communities. For example, practical measures could include repairing deteriorated buildings to enhance safety perceptions and designing more open spaces to alleviate feelings of depression in these areas. By focusing on the needs of vulnerable communities, policymakers can move away from resource allocation strategies based solely on average metrics, leading to more effective improvements in overall well-being. Third, by examining the role of income in shaping sentiment, our research underscores the importance of promoting economic fairness. Income growth is associated with higher predicted sentiment among vulnerable residents and appears to strengthen the positive association between urban visual perceptions and predicted sentiment, whereas this pattern is weaker among wealthier residents. These findings suggest that governments should make more efforts in creating economic opportunities for low-income populations and narrowing income disparities.

5.5 | Limitation and Future Work

A potential limitation of this study concerns the representativeness of sentiment data. Previous studies have shown that social media users tend to exhibit specific demographic characteristics, such as being younger and more digitally engaged (Ilieva and McPhearson 2018). As a result, certain population groups may be underrepresented, particularly residents in low-income MHEs, where individuals are less likely to actively post geolocated content. This sampling bias may lead to an overestimation of residents' sentiment in low-income communities, which could in turn weaken the observed strength of the relationship between urban visual perception and sentiment in the lower MMHI group (Ferrara and Yang 2015).

However, this limitation is partially mitigated in the context of Hong Kong. As one of the most digitally connected cities globally, Hong Kong exhibits high levels of smartphone ownership and internet penetration across socioeconomic groups. Notably, 81.8% of individuals aged 65 and above use the internet, and 96.1% of them participate in online social activities (C&S Department 2023). Thus, while demographic bias cannot be fully eliminated, we believe that the sentiment data still provide meaningful insights. Despite these considerations, an important limitation remains: this study is unable to directly compare the demographics of social media users with census or survey data. Future research could address this limitation by integrating manually collected data, such as household surveys or interviews, with social media data. Such integration would enable a more accurate evaluation of potential demographic bias and help strengthen the interpretability and policy relevance of sentiment-based urban analytics.

Another limitation of this study is that it investigates associations rather than establishing causal relationships between urban visual perception, income, and residents' sentiment. Sentiment is influenced by various unobserved contextual factors, such as social comparison, personal experiences, and environmental conditions (e.g., weather, noise, air pollution). For example, hedonic adaptation and social comparison may partially explain why increases in absolute income do not always correspond to higher happiness. Future research could incorporate these broader environmental and social factors to more precisely capture the relationship between urban visual perceptions and residents' sentiment. Also, applying causal modeling approaches, including causal machine learning techniques with time series data, would help move beyond associative relationships and provide more robust evidence on the determinants of sentiment.

Acknowledgments

This work was supported by the RGC Research Impact Fund (PolyU as PC) (grant number: R5011-23) and the Research Institute for Land and Space (RILS), The Hong Kong Polytechnic University (grant number: CDL1).

Funding

This work was supported by the RGC Research Impact Fund (PolyU as PC) (R5011-23) and Research Institute for Land and Space (RILS), The Hong Kong Polytechnic University (CDL1).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

Alesina, A., R. Di Tella, and R. MacCulloch. 2004. "Inequality and Happiness: Are Europeans and Americans Different?" *Journal of Public Economics* 88, no. 9–10: 2009–2042.

Anselin, L. 1995. "Local Indicators of Spatial Association—Lisa." *Geographical Analysis* 27, no. 2: 93–115.

Benita, F., G. Bansal, and B. Tunçer. 2019. "Public Spaces and Happiness: Evidence From a Large-Scale Field Experiment." *Health & Place* 56: 9–18.

Beute, F., M. R. Marselle, A. Olszewska-Guizzo, et al. 2023. "How Do Different Types and Characteristics of Green Space Impact Mental Health? A Scoping Review." *People and Nature* 5, no. 6: 1839–1876.

Biswas-Diener, R., and E. Diener. 2001. "Making the Best of a Bad Situation: Satisfaction in the Slums of Calcutta." *Social Indicators Research* 55: 329–352.

Blanchflower, D. G., and A. J. Oswald. 2004. "Well-Being Over Time in Britain and the USA." *Journal of Public Economics* 88, no. 7–8: 1359–1386.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32.

Calthorpe, P., and W. B. Fulton. 2001. *The Regional City: Planning for the End of Sprawl*. Island Press.

Camacho-collados, J., K. Rezaee, T. Riahi, et al. 2022. "TweetNLP: Cutting-Edge Natural Language Processing for Social Media." In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, edited by W. Che and E. Shutova, 38–49. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.emnlp-demos.5.

Chang, H., J. Huang, W. Yao, W. Zhao, and L. Li. 2022. "How Do New Transit Stations Affect People's Sentiment and Activity? A Case Study Based on Social Media Data in Hong Kong." *Transport Policy* 120: 139–155.

Chen, T., and C. Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794. Association for Computing Machinery.

Chen, X., Y. Sun, F. I. B. Ibrahim, M. A. B. Kamarazaly, S. N. B. Z. Abidin, and S. Tang. 2025. "Social Media Interaction and Built Environment Effects on Urban Walking Experience: A Machine Learning Analysis of Shanghai Citywalk." *PLoS One* 20, no. 4: e0320951.

Chen, Z., H. Yang, Y. Lin, J. Xie, Y. Xie, and Z. Ding. 2024. "Exploring the Association Between the Built Environment and Positive Sentiments of Tourists in Traditional Villages in Fuzhou, China." *Ecological Informatics* 80: 102465.

Cheng, Z., and R. Smyth. 2015. "Crime Victimization, Neighborhood Safety and Happiness in China." *Economic Modelling* 51: 424–435.

Chunara, R., J. R. Andrews, and J. S. Brownstein. 2012. "Social and News Media Enable Estimation of Epidemiological Patterns Early in the 2010 Haitian Cholera Outbreak." *American Journal of Tropical Medicine and Hygiene* 86, no. 1: 39–45.

C&S Department. 2021. *2021 Population Census—District Profiles*. C&S Department. https://www.census2021.gov.hk/en/district_profiles.html.

C&S Department. 2023. *Information Technology Usage and Penetration*. C&S Department. https://www.censtatd.gov.hk/en/EIndexbySubject.html?scode=590&pcode=C0000031.

C&S Department. 2024. *Table 130–06806: Average Household Size and Median Monthly Household Income of Households by District Council District*. C&S Department. https://www.censtatd.gov.hk/en/web_table.html?id=130-06806.

Diener, E., E. Sandvik, L. Seidlitz, and M. Diener. 1993. "The Relationship Between Income and Subjective Well-Being: Relative or Absolute?" *Social Indicators Research* 28: 195–223.

Diener, E., and M. E. Seligman. 2004. "Beyond Money: Toward an Economy of Well-Being." *Psychological Science in the Public Interest* 5, no. 1: 1–31.

Dolan, P., T. Peasgood, and M. White. 2008. "Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors Associated With Subjective Well-Being." *Journal of Economic Psychology* 29, no. 1: 94–122.

Dolan, R. J. 2002. "Emotion, Cognition, and Behavior." *Science* 298, no. 5596: 1191–1194.

Dong, G., Z. Zhang, H. Zhang, and L. Wu. 2024. "Geographical Match of Objective and Subjective Measures of Well-Being at an Intra-City Scale." *Applied Geography* 167: 103290.

Drukker, M., and J. van Os. 2003. "Mediators of Neighbourhood Socioeconomic Deprivation and Quality of Life." *Social Psychiatry and Psychiatric Epidemiology* 38: 698–706.

Duan, S., Z. Shen, and X. Luo. 2022. "Exploring the Relationship Between Urban Youth Sentiment and the Built Environment Using Machine Learning and Weibo Comments." *International Journal of Environmental Research and Public Health* 19, no. 8: 4794.

Dubey, A., N. Naik, D. Parikh, R. Raskar, and C. A. Hidalgo. 2016. "Deep Learning the City: Quantifying Urban Perception at a Global Scale." In *Computer Vision–ECCV 2016: 14th European Conference*, , 196–212. Springer International Publishing. Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14.

Easterlin, R. A. 1995. “Will Raising the Incomes of All Increase the Happiness of All?” *Journal of Economic Behavior & Organization* 27, no. 1: 35–47.

Easterlin, R. A. 2001. “Income and Happiness: Towards a Unified Theory.” *Economic Journal* 111, no. 473: 465–484.

Easterlin, R. A. 2023. “Why Does Happiness Respond Differently to an Increase vs. Decrease in Income?” *Journal of Economic Behavior & Organization* 209: 200–204.

Ferrara, E., and Z. Yang. 2015. “Quantifying the Effect of Sentiment on Information Diffusion in Social Media.” *PeerJ Computer Science* 1: e26.

Fredrickson, B. L. 1998. “What Good Are Positive Emotions?” *Review of General Psychology* 2, no. 3: 300–319.

Grahn, P., and U. K. Stigsdotter. 2010. “The Relation Between Perceived Sensory Dimensions of Urban Green Space and Stress Restoration.” *Landscape and Urban Planning* 94, no. 3–4: 264–275.

Greenwell, B. M., B. C. Boehmke, and A. J. McCarthy. 2018. “A Simple and Effective Model-Based Variable Importance Measure.” *arXiv* Preprint arXiv:1805.04755.

Guo, X., X. Tu, G. Huang, X. Fang, L. Kong, and J. Wu. 2022. “Urban Greenspace Helps Ameliorate People's Negative Sentiments During the Covid-19 Pandemic: The Case of Beijing.” *Building and Environment* 223: 109449.

Hao, T., H. Chang, S. Liang, et al. 2023. “Heat and Park Attendance: Evidence From “Small Data” and “Big Data” in Hong Kong.” *Building and Environment* 234: 110123.

He, J., J. Zhang, Y. Yao, and X. Li. 2023. “Extracting Human Perceptions From Street View Images for Better Assessing Urban Renewal Potential.” *Cities* 134: 104189.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. “Deep Residual Learning for Image Recognition.” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. IEEE.

Helliwell, J., H. Huang, M. Norton, L. Goff, and S. Wang. 2023. *World Happiness, Trust and Social Connections in Times of Crisis*. World Happiness Report.

Holliday, K. M., A. G. Howard, M. Emch, D. A. Rodríguez, and K. R. Evenson. 2017. “Are Buffers Around Home Representative of Physical Activity Spaces Among Adults?” *Health & Place* 45: 181–188.

Hu, J., S. Ruder, A. Siddhant, G. Neubig, O. Firat, and M. Johnson. 2020. “Xtreme: A Massively Multilingual Multi-Task Benchmark for Evaluating Cross-Lingual Generalisation.” In *International Conference on Machine Learning*, 4411–4421. PMLR.

Ilieva, R. T., and T. McPhearson. 2018. “Social-Media Data for Urban Sustainability.” *Nature Sustainability* 1, no. 10: 553–565.

Ito, K., Y. Kang, Y. Zhang, F. Zhang, and F. Biljecki. 2024. “Understanding Urban Perception With Visual Data: A Systematic Review.” *Cities* 152: 105169.

Jin, T., X. Wei, L. Cheng, K. Wang, Y. Xin, and F. Witlox. 2025. “Street Matters: Linking Perceived Street Environment to Older Adults' Bike-Sharing.” *Travel Behaviour and Society* 41: 101071.

Kahneman, D., and A. Deaton. 2010. “High Income Improves Evaluation of Life but Not Emotional Well-Being.” *Proceedings of the National Academy of Sciences* 107, no. 38: 16489–16493.

Ke, G., Q. Meng, T. Finley, et al. 2017. “LightGBM: A Highly Efficient Gradient Boosting Decision Tree.” In *Advances in Neural Information Processing Systems*, 30. Curran Associates, Inc.

Li, W., R. Sun, H. He, and L. Chen. 2023. “How Does Three-Dimensional Landscape Pattern Affect Urban Residents' Sentiments.” *Cities* 143: 104619.

Lian, A., Y. Zhang, Y. Liu, et al. 2026. “Exploring Sentiment Dynamics and Their Driving Factors in Megacity Residents' Environmental Complaints Through Deep Learning and Multimodal Data.” *Applied Geography* 186: 103806.

Liu, J., K. Shen, X. Liu, and C. Wu. 2025. “Unravelling the Multiple Effects of Multilevel Neighborhood Characteristics on Traffic Crash Risk From a Spatiotemporal Heterogeneity Perspective.” *Travel Behaviour and Society* 40: 101044.

Liu, Y., X. Liu, S. Gao, et al. 2015. “Social Sensing: A New Approach to Understanding Our Socioeconomic Environments.” *Annals of the Association of American Geographers* 105, no. 3: 512–530.

Liu, Z., X. Ma, L. Hu, et al. 2022. “Information in Streetscapes—Research on Visual Perception Information Quantity of Street Space Based on Information Entropy and Machine Learning.” *ISPRS International Journal of Geo-Information* 11, no. 12: 628.

Lou, J., B. Wang, Z. Yuan, and W. Lu. 2024. “Willingness to Pay for Well-Being Housing Attributes Driven by Design Layout: Evidence From Hong Kong.” *Building and Environment* 251: 111227.

Lucas, K. 2012. “Transport and Social Exclusion: Where Are We Now?” *Transport Policy* 20: 105–113.

Lundberg, S. M., and S.-I. Lee. 2017. “A Unified Approach to Interpreting Model Predictions.” In *Advances in Neural Information Processing Systems*, edited by I. Guyon, U. Von Luxburg, S. Bengio, et al., vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

Ma, S., B. Wang, W. Liu, H. Zhou, Y. Wang, and S. Li. 2024. “Assessment of Street Space Quality and Subjective Well-Being Mismatch and Its Impact, Using Multi-Source Big Data.” *Cities* 147: 104797.

Ma, Z. 2023. “Deep Exploration of Street View Features for Identifying Urban Vitality: A Case Study of Qingdao City.” *International Journal of Applied Earth Observation and Geoinformation* 123: 103476.

Mitchell, R. J., E. A. Richardson, N. K. Shortt, and J. R. Pearce. 2015. “Neighborhood Environments and Socioeconomic Inequalities in Mental Well-Being.” *American Journal of Preventive Medicine* 49, no. 1: 80–84.

Mohai, P., D. Pellow, and J. T. Roberts. 2009. “Environmental Justice.” *Annual Review of Environment and Resources* 34, no. 1: 405–430.

Mouratidis, K. 2020. “Neighborhood Characteristics, Neighborhood Satisfaction, and Well-Being: The Links With Neighborhood Deprivation.” *Land Use Policy* 99: 104886.

Mouratidis, K., and A. Yiannakou. 2022. “What Makes Cities Livable? Determinants of Neighborhood Satisfaction and Neighborhood Happiness in Different Contexts.” *Land Use Policy* 112: 105855.

Ochsner, K. N., and E. Phelps. 2007. “Emerging Perspectives on Emotion–Cognition Interactions.” *Trends in Cognitive Sciences* 11, no. 8: 317–318.

Ordonez, V., and T. L. Berg. 2014. “Learning High-Level Judgments of Urban Perception.” In *Computer Vision–ECCV 2014: 13th European Conference on Computer Vision*, , 494–510. Springer International Publishing, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VI 13.

Pessoa, L. 2008. “On the Relationship Between Emotion and Cognition.” *Nature Reviews Neuroscience* 9, no. 2: 148–158.

Pfeiffer, D., and S. Cloutier. 2016. “Planning for Happy Neighborhoods.” *Journal of the American Planning Association* 82, no. 3: 267–279.

Prokhorenkova, L., G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin. 2018. “Catboost: Unbiased Boosting With Categorical Features.” In *Advances in Neural Information Processing Systems*, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, vol. 31. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2018/file/14491b756b3a51daac41c24863285549-Paper.pdf.

Rui, J. 2023. “Exploring the Association Between the Settlement Environment and Residents' Positive Sentiments in Urban Villages and

Formal Settlements in Shenzhen." *Sustainable Cities and Society* 98: 104851.

Salesses, P., K. Schechtner, and C. A. Hidalgo. 2013. "The Collaborative Image of the City: Mapping the Inequality of Urban Perception." *PLoS One* 8, no. 7: e68400.

Sampson, R. J. 2019. "Neighbourhood Effects and Beyond: Explaining the Paradoxes of Inequality in the Changing American Metropolis." *Urban Studies* 56, no. 1: 3–32.

Shaw, M. 2004. "Housing and Public Health." *Annual Review of Public Health* 25, no. 1: 397–418.

Shi, W., M. Goodchild, M. Batty, Q. Li, X. Liu, and A. Zhang. 2022. "Prospective for Urban Informatics." *Urban Informatics* 1, no. 1: 2.

Sinnenberg, L., A. M. Buttenheim, K. Padrez, C. Mancheno, L. Ungar, and R. M. Merchant. 2017. "Twitter as a Tool for Health Research: A Systematic Review." *American Journal of Public Health* 107, no. 1: e1–e8.

Sit, K. Y., W. Y. Chen, K. Y. Ng, K. Koh, and H. Zhang. 2025. "Unveiling Environmental Inequalities in High-Density Asian City: City-Scaled Comparative Analysis of Green Space Coverage Within 10-Minute Walk From Private, Public, and Rural Housing." *Landscape and Urban Planning* 253: 105225.

Smith, K. E., S. Hill, and C. Bambra. 2016. *Health Inequalities: Critical Perspectives*. Oxford University Press.

Tella, R. D., R. J. MacCulloch, and A. J. Oswald. 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics* 85, no. 4: 809–827.

Wang, J., Y. Fan, J. Palacios, et al. 2022. "Global Evidence of Expressed Sentiment Alterations During the Covid-19 Pandemic." *Nature Human Behaviour* 6, no. 3: 349–358.

Winters, J. V., and Y. Li. 2017. "Urbanisation, Natural Amenities and Subjective Well-Being: Evidence From Us Counties." *Urban Studies* 54, no. 8: 1956–1973.

Wu, C., Y. Ye, F. Gao, and X. Ye. 2023. "Using Street View Images to Examine the Association Between Human Perceptions of Locale and Urban Vitality in Shenzhen, China." *Sustainable Cities and Society* 88: 104291.

Yu, Z., K. Shen, and X. Liu. 2026. "Characterizing Walkability in Hong Kong's 15-Minute Transit-Oriented Development (Tod): Insights From Street View Imagery and Local Accessibility." *Travel Behaviour and Society* 42: 101157.

Zhang, F., Z. Fan, Y. Kang, Y. Hu, and C. Ratti. 2021. ""Perception Bias": Deciphering a Mismatch Between Urban Crime and Perception of Safety." *Landscape and Urban Planning* 207: 104003.

Zhang, F., B. Zhou, L. Liu, et al. 2018. "Measuring Human Perceptions of a Large-Scale Urban Region Using Machine Learning." *Landscape and Urban Planning* 180: 148–160.

Zhao, H., J. Shi, X. Qi, X. Wang, and J. Jia. 2017. "Pyramid Scene Parsing Network." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2881–2890. IEEE.

Zheng, S., J. Wang, C. Sun, X. Zhang, and M. E. Kahn. 2019. "Air Pollution Lowers Chinese Urbanites' Expressed Happiness on Social Media." *Nature Human Behaviour* 3, no. 3: 237–243.

Zhou, B., H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba. 2017. "Scene Parsing Through ade20k Dataset." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 633–641. IEEE.

Appendix A

Factor Information

The regression models are developed using 20 control variables encompassing socio-demographic characteristics, built environment attributes, and objective environmental factors, along with six subjective perception variables. Details of control variables, including their descriptions and references, are provided in Table A1. Residents' sentiment serves as the dependent variable. The models are trained on 70% of the data, with the remaining 30% reserved for testing.

TABLE A1 | Neighborhood environmental variables and definitions.

Variables		Description	References
Socio-demographic factor	Proportion of children	The proportion of people who are under 15 years old (%)	Winters and Li (2017); Duan et al. (2022)
	Proportion of elderly	The proportion of people who are aged 65 or above (%)	
	Population density	Number of residential populations per km^2	
	Education level	The proportion of people whose highest level of education is post-secondary (%)	
	Median income	The median of monthly domestic household income (HK$)	
Built environment factor	Building height	The average building height in the community (m)	Benita et al. (2019); Dong et al. (2024)
	Public facilities density	Number of public facilities POIs per km^2	
	Commercial service density	Number of commercial POIs per km^2	
	Leisure service density	Number of leisure service POIs per km^2	
	Educational facility density	Number of education facilities POIs per km^2	
	Financial service density	Number of financial services POIs per km^2	
	POI diversity	Diversity level of POI facilities using Shannon Entropy, C is a set of all POI categories and D_i means density of the ith category: $\text{POI Diversity} = -\sum_{i \in C} D_i ln(D_i)$	
	Street furniture density	The average number of street furniture (e.g., traffic light, street light, pole) observed across all SVIs in the area.	Ma et al. (2024)
	MTR distance	Shortest time to the nearest metro station via path finding algorithm (min)	
Street environment factor	Sea view	The sea view in the SVIs captures elements of the coastal landscape: $\text{Sea View} = P_{water} + P_{sea} + P_{bridge}$	—
	Enclosure	How the street environment encloses pedestrians: $\text{Enclosure} = \frac{P_{building} + P_{tree}}{P_{road} + P_{pavement} + P_{fence}}$	He et al. (2023); Rui (2023)
	Walkability	The overall level of support provided by the outdoor environment for walking: $\text{Walkability} = \frac{P_{pavement}}{P_{road}}$	Ma et al. (2024); Rui (2023)
	Street obstacle	How inaccessible and visually obstructed the street is: $\text{Obstacle} = P_{fence} + P_{railing} + P_{wall} + P_{column}$	Chen et al. (2024); Rui (2023)
	Vitality	Traffic flow reflects the vitality of the street and community: $\text{Vitality} = P_{vehicle} + P_{person}$	Chen et al. (2024)
	Visual complexity	The richness and diversity of visual elements or details within a street scene. We use Shannon Entropy to express diversity: $\text{Complexity} = -\sum_{\cdot \in seg} P_{\cdot} ln(P_{\cdot})$	Liu et al. (2022); Chen et al. (2024)

Note: $P_{\cdot}$ denotes the pixel ratio of $\cdot$ segmentation category. Note, vehicle including car, bus, and truck; pavement includes sidewalk, path, and stair. The pixel ratio is calculated by the proportion of category pixels in total pixel.

Appendix B

Comparison of Regression Models

Table B2 and Figure 4 summarize the results of model training. The "Total" group includes outcomes when the entire dataset is used for training, while "lower MMHI" and "upper MMHI" represent models trained on the lower and upper MMHI groups, respectively. Street environment data is collected from MHEs using buffer zones of 500, 800, and 1000 m. Among the models, CatBoost achieves the best performance, with the lowest MAE, MSE, and the highest pseudo-R^2 across the three groups. The 1000m buffer zone yields the best results, potentially because smaller buffer zones may not fully capture residents' activity ranges.

TABLE B2 | Comparison of performance metrics among different regression methods and buffer zones developed for estimating residents' sentiment.

		Total					Lower MMHI					Upper MMHI				
		Train (Total)		Test			Train (Total)		Test			Train (Total)		Test		
Buffer	Model	MSE	MAE	MSE	MAE	(Pseudo) R^2	MSE	MAE	MSE	MAE	(Pseudo) R^2	MSE	MAE	MSE	MAE	(Pseudo) R^2
500 m	MLR	0.020	0.106	—	—	0.243	0.016	0.099	—	—	0.244	0.020	0.105	—	—	0.331
	RF	0.002	0.032	0.014	0.089	0.428	0.002	0.035	0.021	0.102	0.209	0.002	0.033	0.018	0.099	0.394
	XGB	0.000	0.000	0.013	0.086	0.459	0.000	0.000	0.023	0.110	0.124	0.000	0.000	0.020	0.100	0.340
	LGBM	0.000	0.005	0.012	0.083	0.511	0.000	0.009	0.020	0.102	0.247	0.000	0.008	0.021	0.105	0.298
	CAT	0.000	0.004	0.011	0.080	0.548	0.000	0.001	0.018	0.099	0.313	0.000	0.001	0.016	0.093	0.473
800 m	MLR	0.023	0.114	—	—	0.270	0.017	0.098	—	—	0.284	0.023	0.117	—	—	0.405
	RF	0.002	0.029	0.013	0.074	0.601	0.002	0.030	0.012	0.077	0.436	0.002	0.033	0.023	0.105	0.487
	XGB	0.000	0.001	0.013	0.076	0.608	0.000	0.000	0.013	0.076	0.380	0.000	0.000	0.023	0.107	0.475
	LGBM	0.000	0.006	0.010	0.066	0.707	0.000	0.008	0.015	0.088	0.297	0.000	0.008	0.029	0.119	0.359
	CAT	0.000	0.004	0.009	0.062	0.727	0.000	0.001	0.013	0.076	0.402	0.000	0.001	0.020	0.098	0.543
1000 m	MLR	0.021	0.111	—	—	0.324	0.016	0.095	—	—	0.335	0.020	0.110	—	—	0.473
	RF	0.001	0.024	0.008	0.058	0.768	0.001	0.027	**0.010**	*0.068*	**0.512**	0.002	0.027	*0.021*	*0.098*	*0.528*
	XGB	0.000	0.000	*0.007*	*0.056*	*0.797*	0.000	0.000	0.011	0.070	0.444	0.000	0.000	0.024	0.101	0.467
	LGBM	0.000	0.006	*0.007*	0.062	0.782	0.000	0.007	0.011	0.076	0.422	0.000	0.008	0.024	0.103	0.452
	CAT	0.000	0.002	**0.006**	**0.050**	**0.836**	0.000	0.001	**0.010**	**0.065**	*0.505*	0.000	0.001	**0.018**	**0.089**	**0.595**

Note: Bold numbers indicate the highest rank, while italic numbers represent the second-highest rank in the test results.

Appendix C

The Spatial Distribution of Urban Visual Perceptions

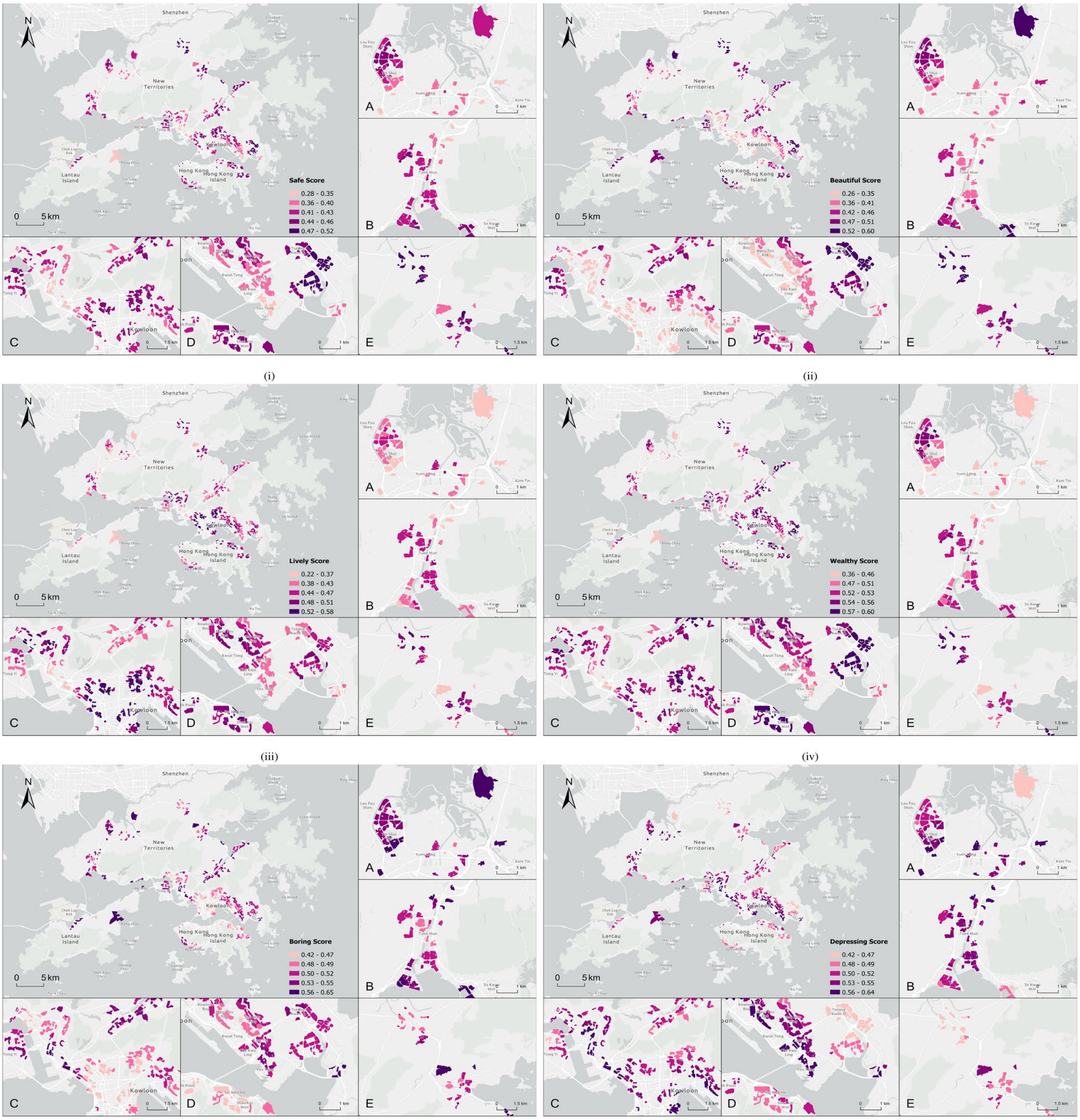

FIGURE C1 | Spatial distribution of urban visual perception scores within MHEs using a 1000 m buffer. Subfigures (i)–(vi) correspond to safe, beautiful, lively, wealthy, boring, and depressing perceptions, respectively. Each subfigure highlights major residential areas in Hong Kong: (A) Tin Shui Wai and Yuen Long, (B) Tuen Mun, (C) Kowloon Area (covering Sham Shui Po, Kowloon City, Wong Tai Sin, and Yau Tsim Mong), (D) Kwun Tong and Sai Kung (including Tseung Kwan O), and (E) Tai Po.