SENSING STREETS: EXPLORING THE ASSOCIATION BETWEEN CITYSCAPE QUALITIES AND STREET PERCEPTIONS USING STREET VIEW IMAGERY AND NATURAL LANGUAGE PROCESSING

SIFAN CHENG[1] and JEROEN VAN AMEIJDE[2]
[1,2]*School of Architecture, The Chinese University of Hong Kong*
[1]*sifancheng@link.cuhk.edu.hk, 0000-0002-2303-7151*
[2]*jeroen.vanameijde@cuhk.edu.hk, 0000-0002-3635-3305*

Abstract. Assessing the perception of street environments and understanding the relationships between their aesthetic qualities and pedestrian experiences are critical to promoting walking behaviour and enhancing urban residents' long-term well-being. While Machine Learning-based analysis of Street View Imagery (SVI) has enabled a range of streetscape studies, the relationship between the visual qualities of cityscapes and people's emotional responses is still under studied. This study used recently developed computational methods to quantify urban street qualities and related sentiments. It collected online reviews and employed Natural Language Processing (NLP) methods to understand how people perceive streets and which environmental features contribute to positive and negative street perceptions. The analytical framework developed in this study can support other high-resolution studies into the spatial-temporal perception of cityscapes in high-density cities across the world.

Keywords. Cityscape quality, street perception, social media data, sentiment analysis, natural language processing

1. Introduction

Understanding the relationships between the aesthetic qualities of cityscapes and pedestrian perception are critical to support urban design aimed at promoting people's walking behaviour and long-term wellbeing. Scholars have worked on conceptual frameworks to link the physical attributes of urban landscapes to environmental perception for several decades (Kaplan, Kaplan & Brown, 1989; Zhang et al., 2018). Most empirical studies used surveys, including interviews and questionnaires, to measure certain dimensions of place-related emotions, and determine people's preferences for certain urban environmental conditions (Roth, 2005; Schroeder & Anderson, 1984). However, due to the time-consuming and labour-intensive nature of traditional survey-based research methods (Roth, 2005), it has been challenging to understand human-environment interactions at a large spatial–temporal scale (Bubalo, van Zanten & Verburg, 2019).

With the development of Information and Communication Technologies (ICTs),

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 2, 139-148.

Street View Imagery (SVI), with its wide coverage, fine spatial sampling, as well as pedestrian perspective, has attracted widespread attention in the assessment of urban cityscape features (Han et al., 2023; Tabatabaie et al., 2023). Meanwhile, social media data are increasingly used to examine associations between urban built environment indicators and people's perceptions or emotional response. Previous studies have applied social media data to explore people's emotions (Mitchell et al., 2013), satisfaction (Lin et al., 2022), and relationship between public sentiment and urban green spaces, public transit and urban renewal (Chang et al., 2022; Wang et al., 2022; Yang et al., 2022). However, the relationship between micro-scale built environment features, especially aesthetic qualities of cityscapes, and people's emotional responses to environments is still understudied.

To address this research gap, this study used recently developed computational methods to sense the urban streetscape. It collected online reviews and employed Natural Language Processing (NLP) methods to understand how people perceive streets and which environmental features contribute to positive and negative street perceptions. Using Sha Tin New Town and Yau Tsing Mong District in Hong Kong as two comparative case study areas, the project quantified public perceptions using sentiment analysis of geo-tagged tweets with two models, VADER and RoBERTa. In parallel, the study obtained Google Street View (GSV) images based on established measurement protocols for street quality (Ewing & Handy, 2009), to quantify the visual qualities of streetscapes. Finally, we assessed the correlation between street quality indices and people's perception indicators.

The research aims to demonstrate the relationships between cityscape aesthetic qualities and people's perception of the built environment, as it explores detailed urban environment studies in a quantitative and human-centric manner. The analytical framework developed in this study can support other high-resolution studies into the spatial-temporal perception of cityscape in high-density cities across the world.

2. Methods

2.1. ANALYTICAL FRAMEWORK

The project used multiple open-sourced urban big data to document the correlation between the cityscape's visual quality and citizen's overall public perception (Figure 1). We employed PSPNet, a machine learning algorithm, to recognize different environmental features of Google Street View images. The presence, visual proportion, and diversity of environmental features were used to quantify the visual quality indices. People's overall sentiment levels were collected from text contents of location-based social network data, and two models, VADER and RoBERTa, were used to predict the sentiment level of English-language geo-tagged tweets. The results of the visual quality and walking perception analyses were plotted into QGIS through associated geographic coordinates. Subsequently, the study conducted statistical analysis on 1) the visual difference between Sha Tin and Yau Tsim Mong; 2) the correlation between visual quality features of cityscapes and people's sentiment-related variables.

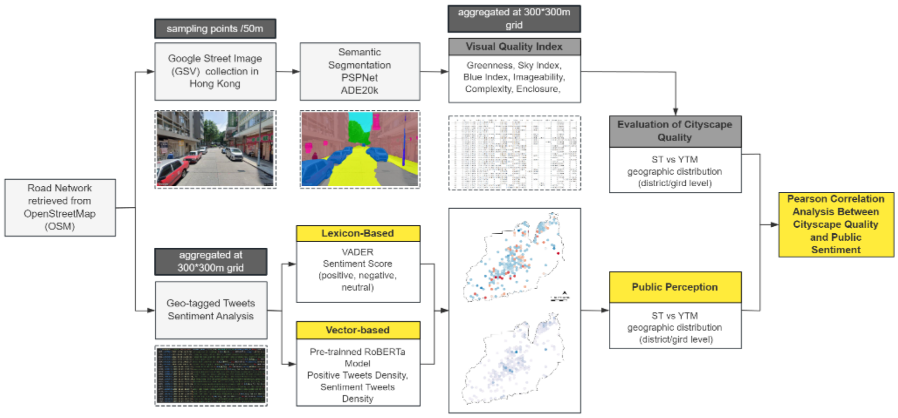

Figure 1. Analytical framework of the study

2.2. CASE STUDY AREAS

Yau Tsim Mong is one of the most famous old urban districts in Hong Kong, containing Mong Kok, Yau Ma Tei and Tsim Sha Tsui. With its unique culture and architectural features, it attracts visitors from across Hong Kong and beyond. Sha Tin is known as one of the most successful new towns in Hong Kong, showcasing a compact and high-density urban layout with multi-level pedestrian connections, balanced land-use mixes and access to waterfront and green areas (van Ameijde. & Cheng, 2021). This study further explores the relationship between urban landscape and levels of public emotion by comparing the results of the two cases.

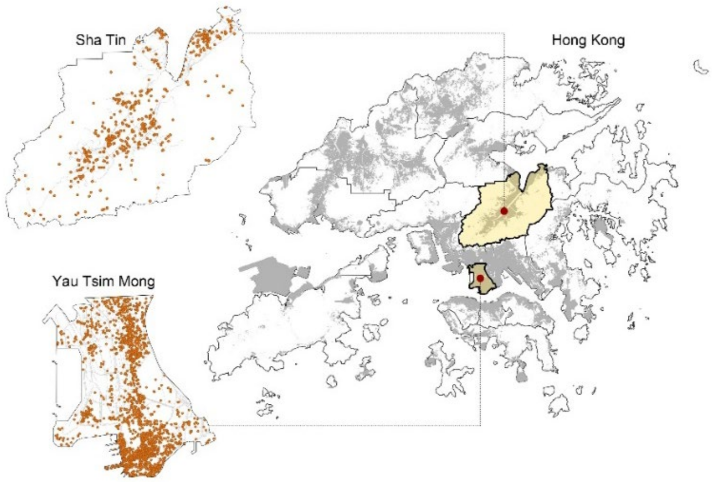

Figure 2. Case study areas. The orange dots indicate the extracted geo-tagged tweets

2.3. DATA GATHERING

We first extracted geospatial road network information from OpenStreetMap. We sampled points every 50 meters, and obtained Google Street View (GSV) images to further assess the cityscape quality. GSV images were requested in an HTTP URL form using the GSV Image API (Google, 2014). For our study, a Python script was created to read the coordinates of each collecting point and evaluate the typological features of the surrounding street networks, to determine the heading view angles for each collecting point. This study used a Semantic Segmentation technique, Pyramid Scene Parsing Network (PSPNet), to extract and compute the view index of different elements. PSPNet has been widely used to evaluate streetscape scenarios due to its high pixel-level precision. We used a pre-trained dataset – ADE20K to process the GSV in our research areas.

For the purpose of evaluating the overall public sentiment, we constructed our Twitter (now named X) database using streaming API and collected tweets posted from May 10, 2019 to May 10, 2023. The collected Twitter data included time, tweet text, user ID, language. After collection, we filtered the data by keeping the geotagged tweets, removing duplicates and tweets from bot accounts. For this study, we kept the English language tweets for our analysis, which accounted for 61% of the dataset. Finally, we obtained 28,369 geocoded tweets posted by 3,954 unique Twitter users.

2.4. OBJECTIVE MEASUREMENT OF PERCEPTUAL QUALITY

Table 1. Visual Quality Indices of Cityscapes

Visual Quality	Significance physical features	Equation
Sky Index	Sky	$Sky_i = VI_{sky}$
Blue Index	Water, Sea, Lake	$Blue_i = VI_{water} + VI_{sea} + VI_{lake}$
Greenness	Tree, Grass, Palm, Plant	$Green_i = VI_{tree} + VI_{grass} + VI_{palm} + VI_{plant}$
Enclosure	Building, Railing, Fences, Ceilings	$Encls_i = \frac{VI_{bldg} + VI_{tree} + VI_{ceil} + VI_{rail}}{VI_{road} + VI_{sidewalk} + VI_{floor}}$
lmageability	Building, Signs (Signboard, Banister, Poster)	$Imgbl_i = VI_{bldg} + VI_{sign}$
Complexity	Person, Building, Benches, Signs, Windowpane, Stage, Poles	$Cmplx_i = \frac{VI_{bldg} + VI_{tree} + VI_{persn} + VI_{benches} + VI_{sign} + VI_{pole} + VI_{win}}{VI_{road} + VI_{sidewalk} + VI_{floor}}$

Subjective perceptions and physical features may have complicated or subtle relations (Ewing et al., 2006; Ewing & Handy, 2009). Initially, Ewing et al. (2006) objectively measured five subjective urban design qualities from the street environment: imageability, enclosure, human size, transparency, and complexity. It has been argued that traditional field observations were limited to a single city, involved a small number of participants, and used low-throughput surveys. Recent studies have employed crowdsourced map data and geotagged images to derive fine-scale perceptual quality information of large urban areas. Ma et al. (2021) and Ye et al. (2019) transformed the perceptual quality indices into the subjective evaluation of a combination of different elements.

To compare the visual quality of varied pilot study areas, we selected four visual quality indices: Greenness, Enclosure, Imageability, Complexity. In addition, previous studies have argued that the nature environment may have strong positive impact on people's sentiment through restorative effects (Kaplan, Kaplan & Brown, 1989), Hence, this study also include the visual index of blue sky as two important indices. The definition and formula for each perceptual quality index is listed in Table 1.

2.5. SENTIMENT ANALYSIS

To decrease the data noise and avoid the 'overfitting problem' in machine learning training, our method first included a series of text data preprocessing by a widely used Python Package for Natural Language Processing, Natural Language Toolkit (NLTK). This involved data preprocessing for English tweets content, including data cleaning, lowercasing, tokenization, lemmatization, and stop words removal.

In this study, we conducted sentiment analysis for each tweet content by two models with different modelling architecture. We first used a lexicon used sentiment analysis model, VADER, to calculate the sentiment score, ranged from -1 to 1. Most of the previous studies employed the widely used sentiment analysis tool—Valence Aware Dictionary for Sentiment Reasoning (VADER; Hutto & Gilbert, 2014). VADER is a lexicon and rule-based sentiment analysis approach that uses natural language processing to combine lexical features with certain grammatical and syntactical rules. However, previous studies have argued that the lexicon-based model has difficulties to understand the in-text content and is sensitive to spelling mistakes or unofficial expressions, which are commonly used in social media content.

To address the limitations by VADER, this study employed a state-of-art sentiment imputation algorithm, BERT (Bidirectional Encoder Representations from Transformers). We used the pre-trained Twitter-roBERTa-base for Sentiment Analysis by Transformers (Loureiro et al., 2022), which is a RoBERTa-base model (Liu et al., 2019) trained on 124M tweets from January 2018 to December 2021, and finetuned for sentiment analysis with the TweetEval benchmark. The model has classified tweet contents into three categories: negative (0), neutral (1), positive (2).

2.6. STATISTICAL ANALYSIS

We first aggregated the result of varied visual quality indices and sentiment score to 300m grids and calculated the average values for all variables. To fully capture the condition of public sentiment by two models, this study involve three aggregated variables at grid level: average sentiment score by VADER model, density of positive sentiment tweets, and density of total sentiment tweets by RoBERTa model. Similarly, we calculate the average value for each grid with four GSV views. A standardized Z-score of each index was then calculated so that the values would be comparable among different indices. We then conducted a descriptive analysis of the Visual Quality indices as well as public sentiment level of 332 geographic grids, to compare the range, mean, and standard deviation of the values. For the three sentiment-related variables, we visualised geographic heatmaps and compare the sentiment level between two cases. Last, we conducted a Pearson correlation analysis, and extracted significantly correlated indicators by evaluating p-values.

3. Results

3.1. THE SPATIAL DIFFERENCES IN VISUAL QUALITY IN YAU TSIM MONG AND SHA TIN

The descriptive results (Table 2) show significant visual quality differences between Yau Tsim Mong and Sha Tin.

The visual greenness of Sha Tin is much higher than in the cityscapes of Yau Tsing Mong. The semantic segmentation result reveals that trees, flowers, palms, and plants were a significant presence in Sha Tin, while only small proportions of trees were found in Yau Tsim Mong. Also, Sha Tin has higher values of blue index, which indicates the higher potential for pedestrians to access rivers and seaside. In contrast, the value of imageability of Yau Tsim Mong is significantly higher than Sha Tin, as a higher proportion of outstanding and unique buildings and skyscrapers can be observed in Yau Tsim Mong area. Surprisingly, the value of complexity in Sha Tin achieved a higher value. This is due to the higher amounts of street furniture, such as benches, signboards, and trees, which were detected. Last, the result shows that over a half of the sampled grids (N =157) achieve a zero value for Blue Index. To avoid errors caused by incomplete data, we removed the blue index in our further correlation analysis.

Table 2. Descriptive Analysis of Visual Qualities

Variable	Statistic	Yau Tsim Mong	Sha Tin
Greenness	Mean	0.1523	0.4452
	Std	0.1395	0.2025
Blue Index	Mean	0.0014	0.0161
	Std	0.0065	0.0788
Sky Index	Mean	0.1156	0.1972
	Std	0.1174	0.1335
Enclosure	Mean	0.0042	0.0201
	Std	0.0285	0.1009
Imageability	Mean	0.6286	0.2473
	Std	0.2213	0.1989
Complexity	Mean	0.0029	0.0154
	Std	0.0202	0.0789

3.2. PUBLIC SENTIMENT DISTRIBUTION

According to the results (Table 3), the sentiment tweets are densely concentrated in

Yau Tsim Mong. The Sha Tin New Town has a much higher proportion of positive tweets according to the RoBERTa model, and higher mean sentiment score in the VADER model. The results generated by the two models indicate that we can observe an overall higher level of public sentiment in Sha Tin New Town. This identifies a potential mismatch between higher-level online activity and sentiment levels retrieved from LBSN data.

Table 3. Descriptive Analysis of Sentiment-related Variables (District Level)

Variable	Statistic	Yau Tsim Mong	Sha Tin
Sentiment tweets counts (RoBERTa)	Positive	1922	391
	Negative	3491	45
	Neutral	21450	1069
	Sum	26863	1505
Sentiment score (VADER)	Mean	0.1198	0.1513
	Std	0.4188	0.3099

Table 4 shows the descriptive results of sentiment variables at grid level. The results show that Sha Tin has a higher average sentiment score (Mean = 0.2088) across grids. They also show the significant difference in sentiment tweets and positive tweets density between Yau Tsim Mong (Mean_pos = 19.375, Mean_sent = 68.095) and Sha Tin (Mean_pos = 0.369, Mean_sent = 1.419). We further investigated the geographic distribution of sentiment variables. Figure 3 shows that the positive tweets density and sentiment tweets density show an overall consistent geographic distribution. Moreover, the sentiment tweets were unevenly distributed, as several grids in Tsim Sha Tsui and Mong Kok achieve extremely high densities of sentiments and positive tweets.

Table 4. Descriptive Analysis of Sentiment-related Variables (Grid Level)

Variable	Statistic	Yau Tsim Mong	Sha Tin
Sentiment score (VADER)	Mean	0.17	0.2088
	Median	0.158	0
	Std	0.1576	0.1195
Sentiment tweets density (RoBERTa)	Mean	68.0946	1.4185
	Median	20.5	0
	Std	121.7369	7.9664
Positive tweets density (RoBERTa)	Mean	19.3784	0.3685
	Median	5	0
	Std	38.1589	4.7756

3.3. SIGNIFICANTLY CORRELATED ATTRIBUTES

This study conducted Pearson correlation analysis among visual quality and sentiment related variables (Figure 4), and we computed the p-values and significant correlated indices. Table 5 indicates that several visual quality indices achieved opposite correlation value in VADER and RoBERTa model.

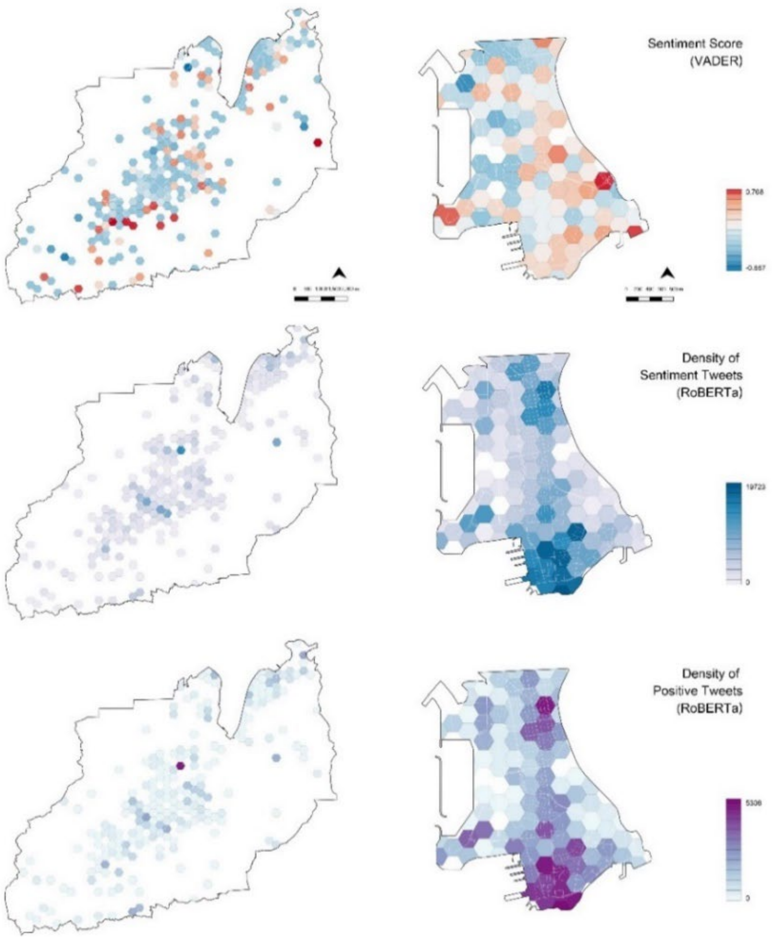

Figure 3. Geographic Distribution of Sentiment-related Variables

For the sentiment scores, greenness and sky index are positively correlated with higher level sentiment (sig.= .214*, .189*). Imageability shows a negative correlation with the sentiment score (sig.= -.275**). The results show an opposite correlation relationship among greenness and imageability, and positive and overall sentiment tweets density. It's noted that greenness is negatively correlated with the density of positive tweets and all sentiment tweets (sig. =-.222*, -.272**), and it also reveals the positive correlation between imageability and positive sentiment tweets density (sig. =.196*, .287**). Also, it can be observed that the density of positive tweets is highly correlated with the density of overall sentiment tweets.

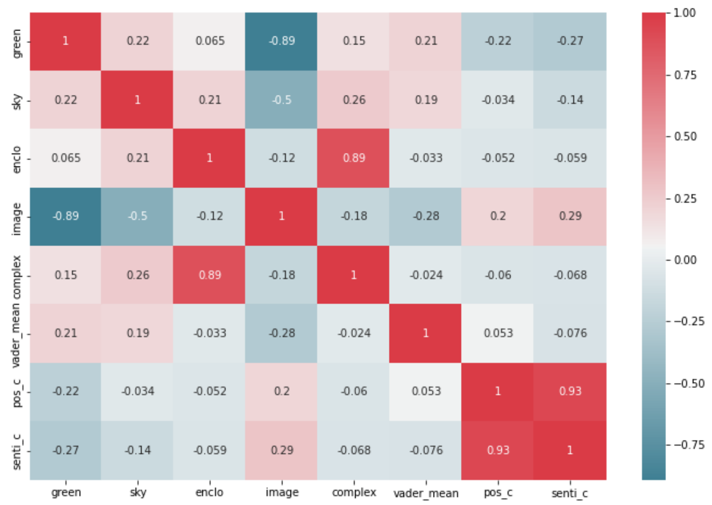

Figure 4. Correlation Analysis Between Subjective/ Objective Street Quality and Perception

Table 5. Significantly Correlated Variables (p-value < .05)

Index1	Index2	Correla-tion	P-value
Greenness	Sentiment score	0.214	0.018
Greenness	Positive tweets	-0.222	0.014
Greenness	Sentiment tweets	-0.272	0.002
Sky index	Sentiment score	0.189	0.037
Imageability	Sentiment score	-0.275	0.002
Imageability	Positive tweets	0.196	0.031
Imageability	Sentiment tweets	0.287	0.001
Positive tweets	Sentiment tweets	0.928	0

4. Conclusions and Future Work

As an innovative empirical study that combines Google Street View index and social media text mining, this study has presented a pilot project for analysing the relationships between cityscape visual quality and citizens' overall public sentiment activities. In short, it can be concluded that (1) there are significant visual quality and sentiment level difference between Yau Tsim Mong and Sha Tin; (2) Yau Tsim Mong shows a higher popularity of online posting activities, while Sha Tin achieves a higher sentiment score; (3) The results of two sets of NLP models shows significant difference in correlation analysis; (4) Several visual quality indices, including greenness, sky index, and imageability are significantly correlated with sentiment-related variables of both models.

In addition, several issues warrant discussion in future work. First, the local variations within the two districts may be studied in further correlation analysis. The mismatch between the density of online posting activity and sentiment levels should be tested across different scales. Second, future studies should explore multiple environment factors which may influence the public sentiment, such as road structures, mixed-use conditions and accessibility of facilities. Last but not least, social media data from various platforms should be integrated to provide a comprehensive perspective of public sentiment.

References

Chang, H., Huang, J., Yao, W., Zhao, W., & Li, L. (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. https://doi.org/10.1016/j.tranpol.2022.03.011

Ewing, R., & Handy, S. (2009). Measuring the Unmeasurable: Urban Design Qualities Related to Walkability. *Journal of Urban Design*, 14(1), 65-84. https://doi.org/10.1080/13574800802451155

Han, Y., Qin, C., Xiao, L., & Ye, Y. (2023). The nonlinear relationships between built environment features and urban street vitality: A data-driven exploration. *Environment and Planning B: Urban Analytics and City Science*, 23998083231172984. https://doi.org/10.1177/23998083231172985

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Kaplan, R., Kaplan, S., & Brown, T. (1989). Environmental Preference: A Comparison of Four Domains of Predictors. *Environment and Behavior*, 21(5), 509–530. https://doi.org/10.1177/0013916589215001

Lin, Y., Wood, S., & Lawler, J. (2022). The relationship between natural environments and subjective well-being as measured by sentiment expressed on Twitter. *Landscape and Urban Planning*, 227. https://doi.org/10.1016/j.landurbplan.2022.104539

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach (arXiv:1907.11692). *arXiv*. http://arxiv.org/abs/1907.11692

Loureiro, D., Barbieri, F., Neves, L., Anke, L. E., & Camacho-Collados, J. (2022). TimeLMs: Diachronic Language Models from Twitter (arXiv:2202.03829). *arXiv*. http://arxiv.org/abs/2202.03829

Roth, M.B. (2005). Online visual landscape assessment using internet survey techniques. Trends in online landscape architecture: Proceedings at Anhalt University of Applied Sciences (pp.121).

Tabatabaie, S., Litt, J. S., & Muller, B. H. F. (2023). Sidewalks, trees and shade matter: A visual landscape assessment approach to understanding people's preferences for walking. *Urban Forestry & Urban Greening*, 84, 127931. https://doi.org/10.1016/j.ufug.2023.127931

van Ameijde, J., & Cheng, S. (2022). Regenerating Hong Kong's New Towns: Resilience and Collaboration in the Context of Polycentric Urban Development. *Urbanie and Urbanus*, Issue 6, P.48 - P.67.

Wang, R., Browning, M. H. E. M., Qin, X., He, J., Wu, W., Yao, Y., & Liu, Y. (2022). Visible green space predicts emotion: Evidence from social media and street view data. *Applied Geography*, 148, 102803. https://doi.org/10.1016/j.apgeog.2022.102803

Yang, L., Marmolejo Duarte, C., & Martí Ciriquián, P. (2022). Quantifying the relationship between public sentiment and urban environment in Barcelona. *Cities*, 130, 103977.