Title	多様な公共空間における感情品質と空間特性の関係に関する研究―マルチタイプデータを用いた機械分類法
Author(s)	李，睿煊
Citation	
Issue Date	2023-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/18404
Rights	
Description	Supervisor: 由井薗隆也，先端科学技術研究科，博士

Doctoral Dissertation

A Study of Relationship between Emotion-Eliciting Quality and Spatial Features in Multi-type Public Spaces—A Machine Classification Approach with Multi-type Data

Li, Ruixuan

Supervisor: Takaya Yuizono

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
[Knowledge science]

March 2023

Abstract

A fundamental problem urban public space designers face is evaluating the spaces having been used for many years. They must develop a reasonable plan that meets the needs of citizens. The classification of emotion elicitation and features of public spaces is an effective method to evaluate the quality of public space and support urban design and decision.

Related studies built several samples of user emotion classification models in public spaces. However, these models' application scope and recognition ability could be improved. In addition, these studies did not answer the quantitative relationship between spatial features and user emotions.

The main goal of this study is to analyze the relationship between users' emotional responses and the features of multi-type public spaces. Furthermore, the main goal was divided into three sub-goals: 1) building an emotion-eliciting quality classification model for multi-types of spaces; 2) extracting the main quantitative features of multi-type public spaces with positive emotional responses; 3) comparing the similarities and differences in the features of public spaces between Japan and China based on users' emotional response.

The study for sub-goal 1 is to build emotion classification models suitable for multi-type spaces using physiological data. To improve the classification accuracy, we chose the ensemble classifiers. The results demonstrate that the highest recognition accuracy of the binary classification model was 94.29%, and the highest accuracy of external validation was 80.90%. In addition, we introduced the synthetic minority oversampling technique (SMOTE) to solve the dataset's problem of too few negative emotion samples. This technology also improved the model's adaptability and met the basic requirements of multi-type public space evaluation.

The study for sub-goal 2 is to extract the main physical and image features of multi-type public spaces for positive emotions. We perform semantic segmentation on spatial photos by introducing a fully convolutional neural network (FCN). Then we obtained the five clusters with different features by two-step cluster analysis. By comparing the value ranges of these spatial features, we got the main spatial features that affect users' emotions.

The study for sub-goal 3 analyzed the similarities and differences in the features between Japan and China by comparing the data on the public spaces' physical, image, and perceptual features. The results show that 1) the differences between Japan and China are more than similarities in the 25 features; 2) the spatial scale, boundary, and continuity of space were the main features that affect the difference between them.

The study results for sub-goal 1 improved the ability of the emotion-eliciting quality classification model, which might contribute to specific urban design practices. The study results for sub-goal 2 found the quantitative features of multi-types of positive spaces, which might be valuable for urban design. And the results of the study for sub-goal 3 explained the similarities and differences in the spatial features between Japan and China from quantitative physical, image, and perceptual features.

In sum, we not only make it clear that there is an association between the features of public space and the emotional response of users but also that different public spaces will have similar results for users. Furthermore, we improved the classification model sample of the emotion-eliciting quality of public space that might be used in practice. We found the quantitative relationship between user emotions and positive spaces, which provides data-based evidence for understanding the relationship between people and space and designing public spaces suitable for human emotions.

Keywords: Multi-type public spaces, Physiological signals, User emotions, Spatial features, Classification models.

Acknowledgments

First and foremost, I am profoundly thankful to my supervisor, Professor Takaya Yuizono, for his guidance during my research. He taught me to find out how to think, how to designate an effective research plan, and how to innovate knowledge. His rich expertise and rigorous scientific attitude constantly put forward new ideas and critical thinking for my research, which brought me into a different field of knowledge science. In addition, over the past four years, he diligently guided my research work in every step, giving me enormous help from purchasing equipment, selecting data collection sites, and formulating research methods for paper writing and publication.

My sincere thanks to my Second Supervisor, Associate Professor Kim Eunyoung, and my minor research advisor Professor Huynh Nam-Van, for their valuable suggestions for my research.

I also would like to thank all the professors who gave me comments or suggestions for my research, especially Professor Kazunori Miyata, Associate Professor Hideomi Gokon in JAIST, and Professor Kiyota Hashimoto from the Graduate School of Prince of Songkla University.

My sincere thanks to all the Yuizono laboratory members who have given me practical suggestions and discussion, especially Wang Ziyu-San, Zhou Shouneng-San, Chigira Yuki-San, Feng Haocheng-San, and all my friends who participated in the field data collection of my research.

Finally, I am grateful to my father, Li Bincheng; my mother, Xin Huixia; my father-in-law Li Lun; and my mother-in-law Tang Shuzhen for their support. I am also very grateful to my wife, Li Xianghui, and my daughter Li Haocong for their love and silent dedication to me so that I can study with peace of mind and complete my original plan and goals.

List of publications

-Papers published in journals

1) Li Ruixuan, Takaya Yuizono, Li Xianghui, Affective Computing of Multi-type Urban Public Spaces to Analyze Emotional Quality Using Ensemble Learning-based Classification of Multi-sensor Data, PLOS ONE (Q1 Journal, IF: 3.58), refereed, 17(6): e0269176, 2022.

-International conference proceedings

1) Li Ruixuan; Takaya Yuizono; Li Xianghui; Wang Yicheng, Exploring the Model of Park Visual Quality Evaluation Using Image Analysis and Machine Learning. In Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering, refereed, pp 333-337, October 24, 2021, Xiamen, Fujian, China.
2) Li Ruixuan, Takaya Yuizono, Huynh Nam-Van, and Li Xianghui, Study on the Relationship between User Emotions and Spatial Characteristics for Public Space Design Decisions: A Case Study of Kanazawa and Nomi in Japan. The 9th International Conference on Civil and Urban Engineering (ICCUE 2022), in press, September 2-5, 2022, Beijing, China.
3) Li Ruixuan, Takaya Yuizono, Li Xianghui, An Approach of Spatial Emotion Recognition Oriented to Urban Design Decision. 2021 International Conference on Education, Management, Economics, Law, and Social Sciences (EMELS 2021), refereed, pp 142-145, May 30, 2021, Hangzhou, Zhejiang, China.

- Domestic conference proceedings

1) Li Ruixuan, Takaya Yuizono, Li Xianghui, An Understanding Method for User's Emotion in Urban Space, New Horizon for Creativity and Innovation—The 41st Annual Conference of Japan Creativity Society, pp 38-41, September 28-29, 2019, JAIST, Nomi, Ishikawa, Japan.

Table of Contents

List of figures

List of tables

Chapter 1
Introduction

1.1 Background

Urban public space is the primary place for citizens' public activities. Due to the diversity of public spaces and the difficulty of emotion measurement, it is challenging to support users' emotional responses in the design and renewal of urban public spaces. So, the emotion-eliciting quality diagnosis and the positive feature extraction of the public spaces based on data-evidence is still one of the topics in urban design.

There are three types of research in this area: the study from the perspective of experts, the study focusing on user experience and evaluation, and the study with the help of sensors, network information, and intelligent technology.

The first type of research obtained the user's behavior information in space through expert observation or taking photos. Experts put forward descriptive suggestions on the design of public space. The conclusions of such studies combined user experience with spatial features, e.g., W. H. Whyte proposed that attractive public space had the features of water, tree, accessibility, and good road conditions (1980). The second type of research focused on users' spatial perception and evaluation. Data acquisition methods include surveys, questionnaires, and interviews. Researchers studied users' needs, emotions, and assessment from the perspective of environmental psychology and proposed main spatial features (Davidson and Milligan, 2004; Weber, Schnier, and Jacobsen, 2008; Gjerde, n.d. 2010; Pallasmaa, 2014; Harvey and Aultman-Hall, 2015; Cho and Kim, 2017). The conclusion of this type of research separated user experience evaluation from spatial features. The third type of research used the data from wearable sensors, media reviews, AI recognition, and other methods to diagnose the spatial quality and extract spatial features.

Among the three types of research, the first two types of research are susceptible to personal factors of researchers or participants because, in some cases, people cannot accurately describe their emotions and may also hide their emotions. These situations make it impossible to identify emotions accurately. The third type of research might reduce the impact of personal factors and external noise by employing a new method of data collection and analysis (Picard, 2000a; Healey and Picard, 2005; Picard, 2010).

The theoretical basis of this research topic is the emotional theories of James Lange (1984), Cannon-Bard (1987), and Schacter Singer (1962), and the cognitive-emotional theory of Lazarus (1991). These emotional theories indicate that emotions are closely associated with environmental stimuli and physiological responses. Environmental stimuli originate from events, weather, people, sounds, images, and landscapes.

Among related research, some researchers studied user emotions in public spaces through physiological reactions. The physiological reactions include external reactions (facial expression, language, and action) and internal physiological reactions (peripheral and central nervous system) (Kreibig, 2010; Kanjo, Al Husain and Chamberlain, 2015). The change in the autonomic nervous system and endocrine activity accompanying the body's emotional response is one way to measure emotion. The physiological signals related to these physiological activities mainly include skin electrical activity (EDA), electromyography (EMG), electrocardiogram (ECG), electroencephalogram (EEG), etc. (Picard, 2000b).

However, there are some limitations in the related research, such as collecting data in a single space, data analysis methods can be improved, and no study on the quantitative features of the multi-type public spaces that elicit positive emotions. Therefore, this study attempts to improve the quality evaluation method of emotion-eliciting quality with the feature extraction of positive spaces.

1.2 Research goals

The main goal of this study is to explore the relationship between multi-types of public spaces and users' emotional responses. This main goal includes three sub-goals.

The first sub-goal is establishing a model of emotion classification in multi-type public spaces based on ensemble learning. This model could diagnose whether a space is positive or negative. The conclusion might help urban managers judge the quality of emotion-eliciting in the public space and decide whether to renew the space.

The second sub-goal is to extract quantitative physical and image features of multiple types of positive spaces. First is to extract useful physical and image features, then to divide public space into high-valence (popular) and low-valence (unpopular) spaces according to emotional evaluation, and finally, to extract the main features of the space. The research results might support the urban design and spatial transformation.

The third sub-goal is to find the similarities and differences in Japanese and Chinese public space' features by comparing physical, image, and perceptual features.

1.3 Research Contribution

This dissertation makes three contributions:

1) **Improving the performance of the public space emotion classification model with physiological data.** Compared with the related research, data were collected in one space, and the model can only classify one space. We collected data in five types of spaces, and the model sample could classify and evaluate emotion-stimulus quality in multiple types of spaces. In addition, the model applied the synthetic minority oversampling technique (SMOTE) to solve the problem of too few negative emotion samples. The results show that the model's binary classification accuracy was 94.29%, and external verification was 80.90%.

2) **The quantitative value range of the features of positive public spaces might improve the urban spatial diagnosis and design.** Using the fully convolutional neural network and unsupervised learning, we obtained the quantitative value range of the spatial features associated with high and low valence. By comparison, we extract data-based evidence of physical and image features of space that affect users' positive emotions.

3) **Finding the similarities and differences between the two countries' public spaces provides a better understanding of the public spaces in Japan and China.**

We used the principal component analysis (PCA) and the entropy weight methods (EWM) to compare the similarities and differences between Japanese and Chinese public spaces in terms of physical features, image features, and perceptual features. The results show that the differences between Japan and China are more than similarities, and the spatial scale, boundary, public/private, and continuity of space were the main features that affect the difference between them. Analyzing the similarities and differences in the two countries' public spaces provides a better understanding of the public spaces in Japan and China.

1.4 Dissertation Structure

This dissertation is organized into eight chapters. After presenting the research background, objectives, and contributions in chapter 1, we discussed the related research in chapter 2 as the basis of this study. Next, chapter 3 describes the research methodology, data collection, and analysis methods. Chapter 4 used physiological signal data to build a multi-type public space emotion classification model. The process includes data preprocessing, feature extraction and reduction, classifier selection, SMOTE processing, model building, model evaluation, and external verification. The purpose of establishing this model is to evaluate the quality of emotional stimulation of space. Chapter 5 described the main extracted features of the spaces with high and low valence using FCN image semantic segmentation and unsupervised learning. These features assist in space transformation, design, and management. Chapter 6 used factor analysis and comparative analysis to compare the similarities and differences in features between Japanese and Chinese public spaces. Chapter 7 concluded this study and discussed the limitation and future works based on the results of chapters 4, 5, and 6. Figure 1.1 below is the content structure of this dissertation.

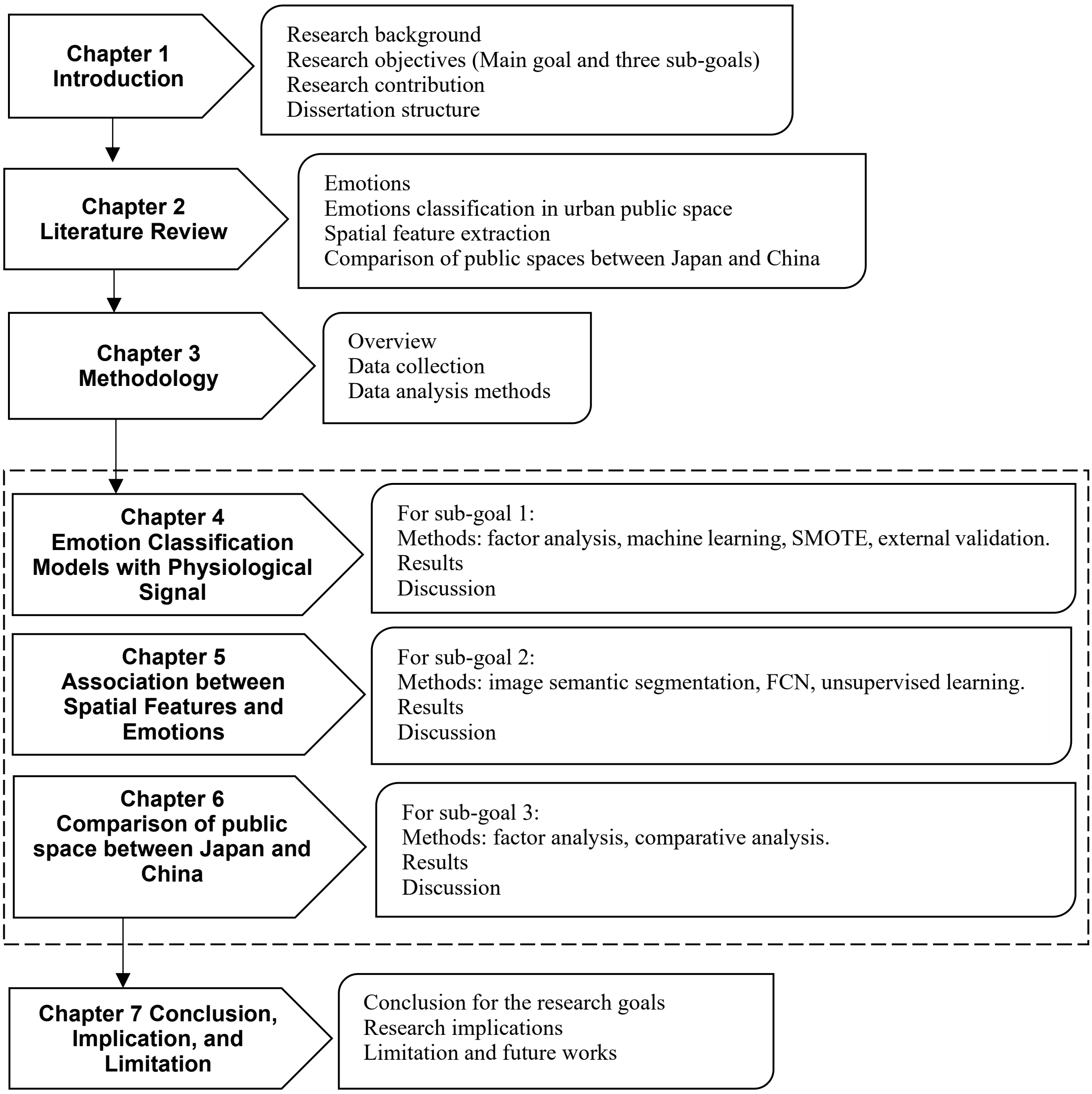


Figure 1.1 Content structure diagram of the dissertation.

Chapter 2
Literature Review

2.1 Emotion

Emotion is a psychological phenomenon of human beings with situational, stimulating, and temporal characteristics. Up to now, the scientific community still has no relatively systematic definition of emotion. Don Hockenbury and Sandra E. Hockenbury proposed that emotion was a complicated psychological state, including three components: a personal experience, physiological response, and behavior or expression response in the book “Discovery psychology” (2014). The Oxford Advanced American Dictionary defines emotion as " a strong feeling such as love, fear, or anger; the part of a person's character that consists of feelings” (Oxford University Press. 2022). Most scholars accept that emotions mainly include three parts: psychological changes, physiological reactions, and external performance. Psychological changes are challenging to monitor, but external performance can be expressed through expressions, sounds, actions, physiological signals, etc.

1) Category of emotions

There are two common classifications of human emotions: basic and dimensional emotions. The former divides human emotions into several basic emotions, and basic emotions could produce complex emotions through mixing (Fischer, Shaver and Carnochan, 1990). The latter suggests that emotions have a multi-dimensional structure and are continuous and gradually changing.

Russell et al. put forward a two-dimensional emotional model in 1979 and proposed two indicators of emotion: valence and arousal (1979). Valence describes the degree of positive and negative emotions; Arousal reflects the intensity of emotions. Figure 2.1 shows the circumplex model of affective proposed by Posner, J., Russell, J.A., and Peterson, B.S.

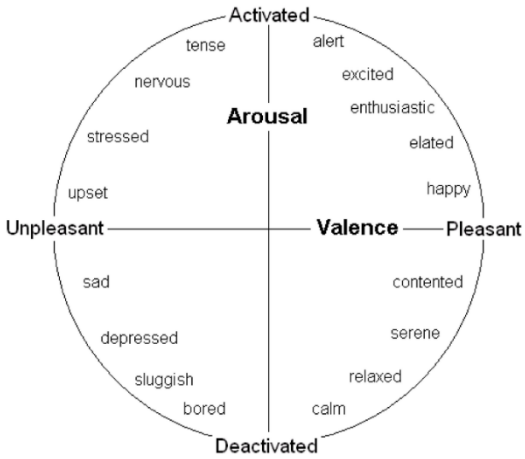


Figure 2.1 The circumplex model of affect (Posner, Russell and Peterson, 2005).

2) Emotion generation

In the 1950s, psychologist M.B. Arnold proposed that the generation of emotion results from the coordinated activities of the cerebral cortex and subcutaneous tissue and put forward the Appraisal Theory of Emotions (Arnold, 1950). This theory states that the emotion generation process consists of three stages: situation stimulating, evaluation, and emotion generation. In the early 1960s, S. Schachter and J. Singer put forward an emotional theory containing two factors through experiments, which refer to the physiological and cognitive arousal of personal emotions (Schachter and Singer, 1962). They emphasized that emotion results from the combination of the surrounding environmental state and the individual's physiological state through the cerebral cortex, which is simply a cognitive process. Various organs in the human body transmit the perceived information to the brain through the perception of environmental factors. This theory was later known as the Schacht Singh theory, and the related model was called the emotional arousal model. Another representative of emotional cognition theory is American psychologist Richard Stanley Lazarus. He proposed that the essence of emotion is an individual's perceptual response to the environmental things around him, resulting from the interaction between people and the environment (Lazarus, 1995). Because the interaction between people and the environment will continue, people will continuously evaluate the surrounding environment on three levels: primary, secondary, and re-evaluation.

Through the above theories about emotional generation, we can summarize that the prerequisite for emotional generation is the influence of external factors, that is, the impact of the environment on people, and personal satisfaction with the environment has a decisive impact on emotion, and also plays a mediating and transforming role between environmental stimuli and emotional responses.

2.2 Emotion assessment method

1) Self-report evaluation method

Self-report assessment methods of emotion mainly include semantic scale and picture-oriented scale. Because semantics can produce ambiguity among people from diverse backgrounds, picture-oriented scales are widely used. Commonly used picture-oriented scales have the day reconstruction method (DRM), the Likert nine-point scale, and the self-assessment manikin (SAM) scale. The SAM designed by Lang (1980) is a widely accepted emotional evaluation method (Figure 2.2). Participants can directly and quickly give feedback on their personal emotional experience when evaluating the sense of pleasure and arousal.

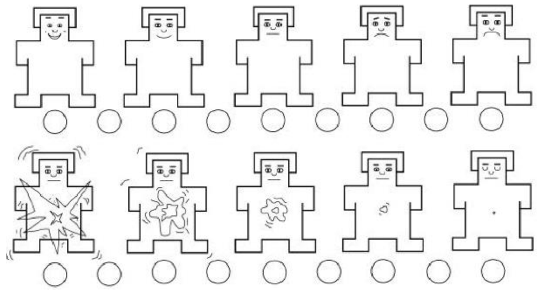


Figure 2.2 Self-Assessment Manikin (Lang, 1980).

2) Physiological measurement method

The physiological measurement method recognizes people’s emotions through the external expression of emotion, such as facial expression, voice, body posture, behavior, physiological signals, etc. (Bradley and Lang, 1994). Physiological signals generally include skin electrical activity (EDA), electromyography (EMG), electrocardiogram

(ECG), electroencephalogram (EEG), respiratory rate, heart rate, blood pressure, skin color, and temperature, which can be observed and measured to obtain the information about emotional changes (Picard, 2000). In these physiological signals, many scholars used EDA, EMG, and ECG for emotion recognition (Alzoubi et al., 2011; Verma and Tiwary, 2014; Torres-Valencia, Álvarez-López and Orozco-Gutiérrez, 2016; Kalimeri and Saitis, 2016; Alberdi, Aztiria and Basarab, 2016; Zhang et al., 2018. p.3886; Nweke et al., 2019; Tan et al., 2020) .

a. Skin electrical reaction (EDA)

Skin electrical reaction refers to the change of skin resistance or skin conductance with the function of skin sweat glands. The human body is excited by the sympathetic nerve, strengthening sweat gland activity, thus secreting more sweat. Because there are more salt components in sweat, the conductivity of the skin is enhanced, forming a skin electrical reaction. The action of the sympathetic nerve causes reactive sweat secretion in human physiological and psychological activities. Skin conductance can be used as an indirect measurement index of the sympathetic nerve. It can also be used as an evaluation index for emotional arousal and psychological activities. Skin conductivity has a fixed resistance parameter. However, skin resistance will decrease under external stimulation or certain emotions, and conductive current will increase (Picard and Scheirer, 2001).

b. Electrocardiogram (ECG)

The Electrocardiogram is a variety of potential change patterns triggered by ECG scanning equipment from the body surface during the successive excitation of a pacemaker, atrium, and ventricle in each cardiac cycle. An electrocardiogram indicates cardiac excitation occurrence, transmission, and recovery. ECG reflects the working condition of the human heart. When the human body is emotional, blood circulation tends to accelerate. With the acceleration of blood circulation, blood pressure, heart rate, and blood vessel volume will also increase.

Heart rate variability refers to the change of heart rate rhythm over time. Heart rate variability is difficult to find through a routine electrocardiogram. The heart rate variability research mainly aims at the time change of each cardiac cycle. Heart rate variability reflects the change in heart rate with different physiological conditions or pathological states (Appelhans and Luecken, 2006).

c. Electromyography (EMG)

Electromyography can judge the functional state of neuro muscle by describing the biological current of neuromuscular unit activity. It is measured by an electrode placed on the skin's surface. Relevant studies argued that muscle tension positively correlates with emotion (Shumailov and Gunes, 2017; Girardi, Lanubile and Novielli, 2017; Hassani et al., 2017). When people are in a positive emotion, their surface muscle tension and voltage will increase. On the contrary, the tension and voltage of surface muscles will decrease under negative emotions. The electrode patches are placed on the face's cheekbones, eyebrows, or arms. Because in the real-world space experiment, if the electrodes are pasted to the participants' cheekbones and eyebrows, their behavior and psychology will be affected. Therefore, the EMG signal can be obtained by measuring the arm and elbow muscle voltage (Egger, Ley and Hanke, 2019).

Most studies state that using multiple physiological signals for emotion recognition is better than using a single physiological signal (Alzoubi et al., 2011; Verma and Tiwary, 2014; Torres-Valencia, Álvarez-López and Orozco-Gutiérrez, 2016; Kalimeri and Saitis, 2016; Alberdi, Aztiria and Basarab, 2016; Zhang et al., 2018. p.3886; Nweke et al., 2019; Tan et al., 2020). Still, they are unclear about the amount of improvement in accurate emotion recognition by using multiple signals.

Features of physiological signals

When physiological signals are used for emotion recognition, extracting features from physiological signals and establishing a recognition model with features as variables is necessary. At present, the main features of physiological signals include time-domain features, frequency-domain features, and nonlinear features (Gong, Ma and Wang, 2016). The time-domain and frequency-domain are the basic properties of signals. Time- and frequency-domain analysis is two observation surfaces of analog signals. Time-domain analysis expresses the relationship between dynamic signals with the time axis as the coordinate; frequency-domain analysis is to change the signal into the coordinate of the frequency axis.

Generally speaking, the representation of the time- domain is more vivid and intuitive. In contrast, the frequency-domain analysis is more concise, and the study of problems is more profound and convenient (Dzedzickis, Kaklauskas and Bucinskas, 2020).

The nonlinear feature refers to the relationship between two variables being not functional, and the image is not a straight line. The joint effect of two factors is only a simple superposition of two independent functions. Typical nonlinear dynamics indicators include the Lyapunov exponent, correlation dimension, approximate entropy, complexity, etc. (Nayak et al., 2018).

Emotion classification based on physiological signals

Picard's pioneering work in 2001 was to apply pattern recognition methods to physiological signal emotion recognition (Picard, 2000). They collected physiological signals from a single subject for several weeks, including electromyography (EMG), blood volume pressure (BVP), skin conductance (SC), and respiration. The results showed that the correct recognition rate of eight emotions was 81%. After that, Picard and Jennifer used the four physiological signals, such as the EMG signal, skin electrical response signal, respiratory signal, and blood volume, to recognize four emotions. A total of 24 statistical features were extracted from the signals, and the final recognition accuracy was 70%; Then, six features were extracted from the physiological signals for feature analysis. The results showed that the recognition rate of anger and calm is 90%-100%, the recognition rate of arousal is 80%-88%, and the emotional recognition rate of pleasure is low, only 50%-82% (Picard, Vyzas and Healey, 2001). In addition, other studies on emotion recognition using different methods have been conducted by scholars in the fields of computer science, psychology, cognition, and physiology in the past two decades (Kim, Bang and Kim, 2004; Wen et al., 2014; Torres-Valencia, Álvarez-López and Orozco-Gutiérrez, 2016; Chen et al., 2017; Ali et al., 2018; Hui and Sherratt, 2018; Al Machot et al., 2019).

In sum, It can be seen that the main contents of emotion recognition research through physiological signals can include the following aspects:

First, most studies stimulate subjects through pictures, music, and video as emotion-inducing media and record the physiological signals of issues in the process of stimulation;

Second, the recorded physiological signals include EEG signals, skin conductance signals, respiratory signals, heart rate signals, ECG signals, EMG signals, and eye movement signals;

Third, for the analysis process of physiological signals, it is generally to extract the physiological signal features at first, then select the features, and finally classify and recognize the physiological features to calculate the accuracy of emotional recognition;

Fourth, commonly used classification algorithms include support vector machines (SVM), neural networks, logistic regression analysis, k-nearest neighbor (KNN) algorithm, decision tree (DT), bayes networks (BN), etc. (Molavi, Yunus and Akbari, 2012; Kanjo, Younis and Sherkat, 2018; Chen et al., 2021a; Keelawat et al., 2021).

2.3 Emotion-eliciting quality in urban public space

Some official institutions and scholars have defined and classified the types of public space. For example, Carr et al. (1992) proposed five public spaces: the downtown park, mini / Vest Pocket Park, pedestrian street, restricted traffic street, and square. Gehl, J., and Gemzøe (2001) divided spaces into five categories according to function: main urban squares, leisure squares, pedestrian streets, traffic squares, and memorial squares. Woolley (2003) divided urban open space into three levels: domestic urban open space, neighborhood urban open space, and urban open space.

Russell and Mehrabian proposed the concept of emotion-eliciting quality in 1978 (Russell and Mehrabian, 1978). They did not define the concept, but we can understand the meaning of it from their article. The various factors stimulate users in the space. When such stimuli generate positive emotions, the space has factors or conditions that generate positive emotions and vice versa. Hence, emotion can be regarded as a comprehensive indicator of the quality of various spatial factors and conditions, that is, the emotion-eliciting quality.

In the fields of psychology, cognition, and computer science, researchers have used various typical or ordinary elicitations (objects) to stimulate participants to elicit physiological responses, collected participants' internal and external physiological data using instruments, and then built models of emotion recognition through data processing and feature extraction (Kreibig, 2010; Ferreira, 2018; Al Machot et al., 2019). In addition, although some researchers have used the same physiological signals as indicators, the features and classifiers were different (Nasoz et al., 2004; Geiser and Walla, 2011; Verma and Tiwary, 2014; Torres-Valencia, Álvarez-López and Orozco-Gutiérrez, 2016; Ali et al., 2018; Zhang et al., 2018; Kołakowska, Szwoch and Szwoch, 2020; Li, Zhang and Song, 2021). Different signal features lead to no comparability of research results. Therefore, it is imperative to screen the main features and compare the classifiers to obtain a more reliable evaluation model.

Some urban design and geography researchers have introduced emotion recognition methods and conducted related experiments in urban spaces (Table 2.1). They selected a single type of space for data collection, such as a predefined route in the city center (Birenboim et al., 2019), a shopping route in a city center (Kanjo, Younis and Sherkat, 2018), a specific route around a city center (Kanjo, Younis and Ang, 2019), or predetermined route in a neighborhood (Ojha et al., 2019), which makes data collection easy. However, this approach limits the scope of the application of the model. Besides, researchers mainly used six single classifiers: single-classification support vector machine (SVM), k-nearest neighbors (KNN), naïve Bayes (NB), convolutional neural network, long short-term memory (CNN-LSTM), multilayer perceptron (MLP), and one ensemble classifier random forest (RF), and finally developed binary, ternary, and quinary emotional classification models.

Table 2.1 The related studies on emotion recognition using physiological sensors in urban spaces over the decade.

Reference	Sites	Number of subjects	Signals	Number of features	Emotions	Classifier	Results
Olsen and Torresen (2016)	a natural environment of daily life	10	accelerometer data	8 features	valence, arousal (3 classes)	SVM	arousal 75%, Valence 50.9%
Kalimeri and Saitis (2016)	a selected route in a campus	9	EEG, EDA	182 EEG features, 6 EDA features	5 predefined categories	RF	79.30%
Kanjo et al. (2018)	a shopping route in a city center	40	EDA, HR, and body temperature	21 features	emotions	SVM, RF, KNN, and NB	86%
Birenboim et al. (2019)	a predefined route in the city center	15	EDA, HR, and HRV	5 EDA and 3 HRV features	stress level	N/A	N/A
Ojha et al. (2019)	a predetermined route in a neighborhood	30	EDA	9 features	valence and arousal	REP-Tree, MLP, and SVM	87% (binary), 80% (multi-class)

Up to now, emotional classification research based on physiological signals generally includes the following steps:

a. Selecting physiological signal feedback instruments and related equipment;

b. Selecting emotional stimuli;

c. Conducting experiments and collecting physiological signals;

d. Extracting and reducing signal features;

e. Selecting classifiers,

f. Building model;

g. Model evaluation.

2.4 Spatial feature extraction

1) Spatial features related to emotions

Some researchers have extracted spatial indicators and identified relationships between spatial features and user emotions. Table 2.2 shows the related research conducted since 2006. Lee et al. (2009) investigated pedestrians' preferences for sidewalks and proposed main indexes affecting sidewalk comfort and user emotions: sidewalk width, shrub width, tree height, and treewidth. Ewing, R. and Handy, S. (2009) applied the expert group rating method. They proposed that five types of urban design quality can be measured according to the physical indexes of streets: imageability, closeness, human scale, transparency, and complexity. Based on the principle of Kansei engineering, Xiang, L. and Papastefanou, G. (2019) studied the correlation between the urban environment's physical features and pedestrians' emotional responses. They found that open space with visual objects in the distance is the main factor in generating positive emotions.

In contrast, a closed refuge space is not a positive factor of happiness, contradicting the results of previous studies. Schneider et al. (2014) used street-view images and virtual reality (VR) systems to realize immersive virtual space to test participants' emotional reactions. Testers controlled this technology through physical interfaces to change the five street parameters: street width, building height, building width, the distance between buildings, and the number of buildings. This method makes it easier to determine the relationship between spatial features and emotions; however, the research is still in the preliminary testing stage.

Some researchers put forward spatial features that affect user emotions; some combine the physical, aesthetic, and functional features of urban space. For example, Carmona et al. (2010) combined the functional indexes on “activities and social communication” with the physical indexes on “sunlight, shade, fountains, air pollution, and wind movement” without distinguishing their weight. Kalivoda et al. (2014) did not indicate the different correlation degrees between physical indexes such as “vegetation” and aesthetic indexes such as “style unity.” Zhang et al. (2018) regarded the weights of aesthetic indexes as “lively, boring, wealthy, depressing, and beautiful” and that of physical indexes as “wall, fence, field, hill, bridge, pole, sidewalk car, vegetation, landmark, transportation, and water body” as the same.

Furthermore, other researchers selected indexes that are not comprehensive. For example, Lee et al. (2009) analyzed physical indexes such as ‘sidewalk width, shrub width of sidewalks, tree height, tree width, green ratio, sky ratio, roadway ratio, sidewalk ratio, and building ratio,’ which are mainly scale elements or components; however, these indexes lack spatial boundary elements and spatial continuity elements. The physical indexes proposed by Cho, M. E. and Kim, M. J. (2017) combined aesthetic indexes of ‘complexity, variety, complexity, and variety with physical indexes of ‘facade, material, color, texture, and light’ but did not discuss other types of spatial indexes.

Table 2.2 Related studies on spatial feature extraction in urban public spaces.

Year	Authors	Measures and techniques	Number of participants	Evaluation indexes
2006	Borton & Mitchell	Questionnaire	N/A	Simplicity, harmony, clarity of visual perception, comfort, security
2009	Ewing et al.	Video clips of streetscapes, questionnaire	10	Proportion of historic buildings, outdoor dining, major landscape features, proportion of street walls, long sight lines
2009	Lee et al.	Questionnaire	102	Sidewalk width, shrub width, tree height, tree width, green ratio, sky ratio, roadway ratio, sidewalk ratio, building ratio
2010	Carmona et al.	Questionnaire	N/A	Movement and activity, organising the movement of vehicles and pedestrians, communication and optical permeability, activities, social communication, privacy, population density, open public space, benefit from the infrastructure of the place, sunlight and shade, fountains, air pollution, wind movement
2014	Kalivoda et al.	Questionnaire, pictures	400	Vegetation, style unity, proportion of uniformity, symmetry
2014	Schneider et al.	VR headset, a game engine, physical interface, self-reporting	N/A	Width of the street, height of buildings, width of buildings, distance between buildings, number of buildings
2016	Mamaghani et al.	Interview, questionnaire, SD method	80	Pleasant, exciting, relaxing, distressing, coherent, natural, complex
2016	Jahanmohan, T.	Photography survey, mind maps	author	Courtyards, plazas & parks, major landscape features, proportion of historic buildings, buildings with identifiers, buildings with non-rectangular shapes, outdoor dining, noise level, long sight lines, street wall, sky, street trees, street furniture, building height, small planters, windows at street level, active uses, buildings, building colours, buildings accent colours, public art, pedestrians
2016	Li et al.	Wristband sensor, GIS	30	Building shapes and textures, isovist parameters, visual entropy, visual fractals, enclosed urban spaces, landscape hierarchy, greenery
2016	Ernawati et al.	Questionnaire	103	Legibility, human scale, coherence, imageability
2017	Cho M E et al.	Self-report, questionnaire, conversation, physiological sensors	3	Volume and style, facade, complexity, variety, proportion, rhythm, unity, material, colour, texture, light
2017	Liu et al.	Google Street View, machine learning	752	Quality of building facade, visual continuity of street, architectural style, scale of the building, relationship between street and adjacent buildings
2018	Bivina et al.	Questionnaire	2804	Sidewalk surface, sidewalk width, obstruction, the potential for vehicular conflict, continuity, encroachment, availability of crossing facilities, security, walking environment, comfort
2018	Fathullah & Willis	Physiological sensors-EDA	9	Green elements, natural characteristics, traffic level
2018	Zhang et al.	Street View pictures, machine learning	N/A	Safe, lively, boring, wealthy, depressing, beautiful. Wall, fence, field, hill, bridge, pole, sidewalk car, vegetation, landmark, transportation, waterbody
2018	Ernawati & Sudarmo	Questionnaire, Likert scale	110	Enclosure, legibility, human scale, transparency, complexity, coherence, linkage, imageability
2019	Tang et al.	Tencent Street View, machine learning	N/A	Street facade colour, facade materials and decoration, parking, greenery, street furniture, sign boards, store facade transparency, green space, the degree of openness, enclosure, continuity of wall street, width to height ratio of street section
2019	Xiang & Papastefanou	Portable smart wristband	30	Open space with visual object in the distance, enclosed place
2019	Steinmetz-Wood et al.	Google Street View, questionnaire	136	Sidewalk continuity, sidewalk buffer, sidewalk quality, benches, streetlights, curb cut quality, traffic calming, street traffic lights, building setback, building design variation, presence of trees, shade, nature areas, landscaping, landscape maintenance

The results of the Related studies have shown that images can stimulate people to produce different psychological responses (Zhao, 2016; Zhao et al., 2018). The low-level visual features of an image, such as color, texture, and shape features, can only analyze the visual information of the image, and there is a problem of insufficient information (Zhang et al., 2011; Peng et al., 2015; Priya and Udayan, 2020). In 1996, Eakins put forward the term "image emotional semantics" at the third international e-book and visual information retrieval conference (1996). He suggested that the image content semantics should be the primary carrier of image emotional information, and the content level information is closer to the perceptual features of people when they observe images. With the application of deep learning technology, some researchers began to use the convolutional neural network algorithm for image semantic segmentation, that is, the segmentation of object elements in the image and the correlation between object features and visual evaluation in the image (Wang and Wang, 2005; You et al., 2016; Peng et al., 2016).

From the results of related research, we found that most researchers agreed that space's physical and image information features affect users' emotions. However, these studies lack interconnection, resulting in no comparability and possibility of integration (Table 2.2).

2) Methods of spatial feature extraction

There are generally three methods for the extraction of spatial features. First, most related studies used the main spatial features selected by the researchers according to their professional background (Harvey, C, 2014 Harsritanto, Indriastjario and Wijayanti, 2017; Hooi and Pojani, 2019; Oliveira, 2022;). Some researchers also used questionnaires and conversations to ask participants to select the features they think are important from the features determined by the researchers (Peschardt and Stigsdotter, 2013; Jiang, Chang and Sullivan, 2014). The second method proposes features that affect spatial quality through participant self-reports. This method includes two ways: one is that the participants give feedback after walking in the real space; the other is that the participants give feedback by looking at pictures or using VR to visit the virtual environment. The third method is that some studies utilized image semantic segmentation techniques to extract spatial image features. For example, Vilera, R., Rachmadi, R.F. and Yuniarno, E.M. (2020) proposed segmentation and selective feature extraction of street view images captured during the day using semantic segmentation technology, which has strong robustness and a better understanding of each object and occluded objects in street view images.

The first method mentioned above is proposed by professionals, which can fully reflect the information of the space. Still, if the participants have no professional background, it will cause them to misunderstand the features, and they will not get accurate feedback information. The second approach was to extract spatial features from participants' self-reports. This method can directly obtain the experience feedback of the participants, but the description of the spatial features may be vague, and it is not easy to obtain quantitative data. The third method can extract the proportion of objects in the spatial image, which belongs to the spatial image features. However, this method ignores the position and layout of things in space, so the obtained spatial information is insufficient.

2.5 Comparison of public spaces between Japan and China

In urban design, some comparative studies are on Japanese and Chinese traditional gardens and their cultural backgrounds (Zhao et al., 2003; Chen et al., 2021b; Han, 2022). Zhao et al. (2003) assessed the landscape images of Japan and China and analyzed the factors that affect these images. Then through psychological investigation, it is found that there are many differences in the image display of the themes of "Japanese style" and "Chinese style," and the image of urban public space is listed. Han, Y. (2022) compared the Zhanyuan Garden in Nanjing and the Longan Temple in Kyoto, analyzed the cultural background of these gardens, and proposed the similarities and differences in the philosophical systems, aesthetic perspectives, and landscape techniques that formed these gardens. Chen et al. (2021b) applied principal component analysis to compare the traditional gardens in Kyoto and Suzhou and the modern gardens in Xiamen and try to find the similarities and differences between the three garden environments.

These comparisons are scattered, and there is no systematic analysis, and most of them are qualitative studies, which are challenging to support the urban design and municipal management directly. Therefore, we compare the similarities and differences between urban spaces in Japan and China from the quantitative features of more specific spaces and provide an understanding of the public spaces between the two countries.

2.6 Research gaps in the existing literature

Through the above statements of previous relevant studies (Tables 2.1 and 2.2), we found that:

1) Some previous studies on classifying spatial emotions based on physiological signals were conducted in the laboratory, and a few were conducted in the real world. Since the virtual environment removes factors such as space noise, people, and weather, it is not easy to obtain the actual emotional response of people in the real world. The model established by this research has little practical significance.

2) Because the urban public space is diverse and complex, but the previous related research is to collect data in one kind of urban space, the model built through these data lacks scalability.

3) Previous studies lack the step of feature reduction. It is necessary to reduce the number of features to reduce the degradation of model performance caused by redundant information. Moreover, related studies did not deal with the imbalance of sample size, resulting in problems in the recognition performance of the built model. Urban public space is a weak emotional stimulus for people. There are fewer very high and low valence cases, so it is necessary to balance the data set. And previous studies mainly used single classifiers and fewer ensemble classifiers. It is possible to improve the recognition ability of the models by ensemble classifiers.

6) Related studies have built the spatial emotion recognition model and extracted spatial features, but the two have not been combined. Because the emotional model is used to evaluate the emotion-eliciting quality of space, it does not tell urban designers

what the spatial features of high emotion-eliciting quality are. What are the spatial features of low emotion-eliciting quality? Therefore, extracting quantitative features of positive space may further develop and refine the evaluation model results.

2.7 Summary

This chapter conducted a literature review of related research on the emotion classification model of urban public space based on physiological signals, the extraction of spatial features, and the comparison of public space features between Japan and China. We observed that data are collected from a single space in the related research on the public space emotion classification model based on physiological signals. A single classifier is mainly used to build the classification model. Due to the particularity of the space, the scope of application of the models is limited. Relevant studies mainly use the methods of researcher classification or questionnaires and participant self-reports to extract spatial features. There is a problem with the mixed use of different types of features. In the comparative study of the features of public space between Japan and China, the related research mainly conducts a comparative analysis of the garden environment of Japan and China from the aspects of cultural background and aesthetic preference. It lacks the comparative study of public space at the feature level. Faced with these problems, we will try to propose new methods, introduce new technologies, and try to make complete and practical research.

Chapter 3
Methodology

3.1 Overview

The objects of this research are urban public spaces and users. The main goal is to study the relationship between users' emotions and public space features. We divided the main goal into three sub-goals. The first sub-goal is to build a model to identify the spatial emotion-eliciting quality in public spaces. We used physiological signals and SAM self-reported data to build the recognition model based on the user's physiological signals and user emotions. This model is used to identify whether an urban public space is positive or negative (binary classification), positive, neutral, or negative (ternary classification), or a more refined five-class classification. The second sub-goal is to extract the positive space's physical and image features and find the quantitative indicators of the main features that affect users' emotions. The third sub-goal compares the similarities and differences in public spaces between Japan and China regarding physical, image, and perceptual features.

During the data collection process, we collected five types of data:

1) Data from the SAM scale (questionnaire);

2) Physiological signal data (EDA, ECG, EMG).

3) Spatial perception evaluation data (questionnaire);

4) Physical feature data (measured in real space);

5) Photos of space sections (shot along paths in public spaces).

Both SAM scale data and physiological signal data (EDA, ECG, EMG) are associated with emotions in the five kinds of data. We might get emotional feedback directly from the SAM, but there is no information about the features of emotion required by machine learning; Physiological signals can not tell the emotion feedback directly but can tell the features related to emotions. So, we combined them to build the spatial emotion-eliciting quality classification model. The impact of SAM data on the model is weakened by training with a large number of samples. In practical application, instead of collecting SAM scale data, we only collect physiological signals, input their features into the built model, and obtain the evaluation results of the spatial emotion-eliciting quality.

After obtaining the five types of data, we used different techniques to process and analyze these data; the following are the data and skills we use in target research:

1) Sub-goal 1: building a spatial emotion classification model.

Data: data from SAM scale, physiological signal data.

Analysis methods: principal component analysis (PCA), ensemble learning, synthetic minority oversampling technique (SMOTE).

2) Sub-goal 2: extraction of spatial features.

Data: data from SAM scale, physical feature data of spaces, photos of space sections.

Analysis methods: image semantic segmentation based on a fully convolutional neural network (FCN), Two-step clustering algorithm, and categorical principal components analysis (CATPCA).

3) Sub-goal 3: comparison of public space between Japan and China.

Data: physical feature data of spaces, photos of space sections, spatial perception evaluation data.

Analysis methods: PCA, the Mann-Whitne U test, and the Chi-square test.

Figure 3.1 below shows the research framework.

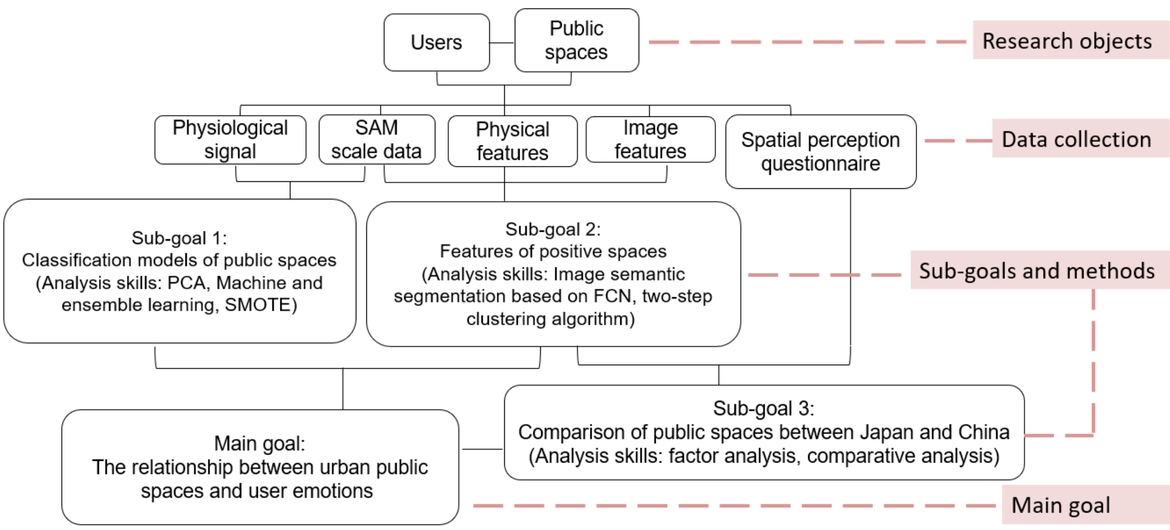


Figure 3.1 Diagram of the research framework.

3.2 Data collection

According to the Regulations on the Conduct of Research Involving Human Subjects of the Japan Advanced Institute of Science and Technology (JAIST), we submitted a human body research plan to the Research Ethics Committee of JAIST and obtained research permission before the data collection. The research process followed the principles of the Declaration of Helsinki. All the participants signed the informed consent.

3.2.1 Site selection

This study conducted an on-site experiment in urban spaces with different functions to obtain general results. We collected data from 10 public spaces of five types: five in Nomi City, Kanazawa City, Japan, and five in Dalian City, China. The five types of spaces were campus public spaces, residential areas, park spaces, memorial spaces, and historical pedestrian street spaces.

There are two reasons why we chose these five types of spaces; one is that these five space types are the main types of public space. The other is that it is an attempt for us. The related studies have chosen one type of space to collect data, and more space types might have too much data and unpredictable results.

The reason why we chose the location in Nomi city, Kanazawa city, and Dalian city is that these locations are closer to JAIST or Dalian Polytechnic University, it is convenient to collect data, and there is also consideration of costs (participants' fee, transportation fee, and lunch fee, etc.). It took ten days to collect data, including five days in Japan and five days in China.

We selected a linear space of approximately 300–1000m as the experimental route in each space. We divided each route into four sections with different spatial features (function and structure) for $10 \times 4 = 40$ sections. The reason why we divided the route of each space into four sections is that the attributes of these four sections are different so that users may generate different emotional responses (positive or negative) in each section, and it will be difficult for users to fill the SAM if they are not divided.

Additionally, we divided these ten spaces into the ratio of 8:2. We used the data from eight spaces for model training and testing and the data from the other two spaces for the external validation of the built model. The location, function, sections, and length of the selected spaces and experimental routes are listed in Tables 3.1 and 3.2. Figures 3.2 and 3.3 show the route maps and photos of each section. We used the data from the spaces in Figure 3.2 and Table 3.1 to train and test the models and those in Figure 3.3 and Table 3.2 to verify the model performance through external validation.

Table 3.1 Basic information about eight spaces where data was collected for model training and testing.

Sites	Latitude and longitude	City	Function	Total length (m)	Section 1 (m)	Section 2 (m)	Section 3 (m)	Section 4 (m)
JAIST campus	36.4444, 136.5924	Nomi city (Japan)	Campus	697.5	180.4	123.8	192.3	201
Residential area in Yokaichi	36.5463, 136.6088	Kanazawa city (Japan)	Residential area	1067.9	188.1	280.6	299.3	299.9
D. T. Suzuki Museum	36.5578, 136.6608	Kanazawa city (Japan)	Monument	364.7	77.3	79.1	164.7	43.6
Higashi Chaya District	36.5720, 136.6668	Kanazawa city (Japan)	Pedestrian street	649.4	100.4	241.6	204.6	102.8
DLPU campus	38.9713, 121.5278	Dalian City (China)	Campus	651	161	136	251	103
Meilin Park residential area	38.9734, 121.5177	Dalian city (China)	Residential area	945	285	262	174	224
Huarun 24 City Park	38.9755, 121.5365	Dalian city (China)	City park	719	114	86	320	199
Dalian Heroes Memorial Park	38.9007, 121.6234	Dalian city (China)	Memorial site	788	247	95	294	152

Table 3.2 Basic information about the two spaces and their data were used for model external validation.

Sites	Latitude and longitude	City	Function	Total length (m)	Section 1 (m)	Section 2 (m)	Section 3 (m)	Section 4 (m)
Kenroku-en	36.5622, 136.6626	Kanazawa city (Japan)	Japanese garden	668.8	208.4	116.9	173.2	170.3
Dalian Ganjingli · Dongshi historic district	38.9508, 121.5350	Dalian city (China)	Pedestrian street	621.0	149.0	144.0	206.0	122.0

3.2.2 Data collection in public spaces

A total of 20 Chinese students (7 men and 13 women; average age, 28.6. Fourteen were aged 20–29, four were aged 30–39, and two were 40–49) participated in the experiment. Nine experiments were conducted in Nomi City and Kanazawa City, Japan, and 11 participated in the investigation in Dalian City, China. The dates of the investigation are October 2019 and October 2020, respectively.

Figure 3.2 Route maps and photos of each section in the eight spaces where data collected were used for model training and testing.

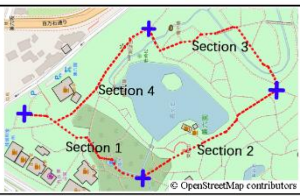
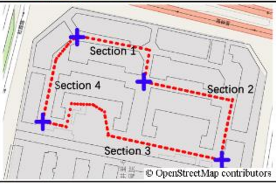
Urban space	Route map	Photo of section 1	Photo of section 2	Photo of section 3	Photo of section 4
Kenroku-en					
Dalian Ganjingli· Dongshi historic district					

Figure 3.3 Route maps and photos of each section in the two spaces where data collected were used for model external validation.

Except for the two campuses, none of the participants visited any sites before the data collection. Before data collection, the aims and experiment content were explained to each participant. All the participants signed a formal consent form. We chose to collect data on sunny days. During the data collection, the participants wore a Bitalino portable physiological signal feedback instrument (BITalino (r)evolution Plugged kit, PLUX Wireless Biosignals Ltd., Portugal) (Figure 3.4), and a head-wearing DV(Ordro), carried a GPS device (Nav-uNV-U73T, Sony) and a laptop (Surface), and walked through the spaces. The physiological signal feedback instrument collected the participants' EDA, ECG, and EMG, which were stored on a laptop in the backpack through the supporting software OpenSignals_(r)evolution_win64 (Figure 3.5) (Among all the physiological signals, EDA, ECG, and EMG are mainly used in related studies (Table 2.1), these three physiological signals may be reliable for these studies. So, we chose the physiological signals of EDA, ECG, and EMG).

The GPS recorded the participants' location information simultaneously. Each participant filled out the SAM scale (Figure 3.6) and the spatial evaluation questionnaire (Table 3.3) immediately after walking through each space (Figures 3.7 and 3.8). In the spatial evaluation questionnaire, 12 pairs of antonym adjectives were used to describe the spatial perception of users. Five semantic levels were designated for each pair.

Meanwhile, the spatial features extracted in the ten public spaces were divided into two types: the spatial interface feature and the proportion of objects in the spatial image (walking in the middle of the set path). The data of interface features were obtained through on-the-spot measurement and Google Earth. Twenty-four physical space features of six items were extracted from each section's front, middle and back sections. The six main elements include scale, boundary, spatial continuity, visual, thematic, and components. The 24 features could be quantified by direct measurement or grading methods.

In addition, before the data collection, we chose to take photos of all the routes between 9 AM and 4 PM on a sunny day in October. Walking along the middle of the routes of the public spaces, which is consistent with the route of the participants, we took a photo every ten steps (about 5.5m) from the starting point, and the camera shooting direction was parallel to the route. The photographer's height is 1.72m, so the camera is about 1.62m from the ground. After shooting at all sites, we got 1172 photos (503 in Japan and 669 in China).

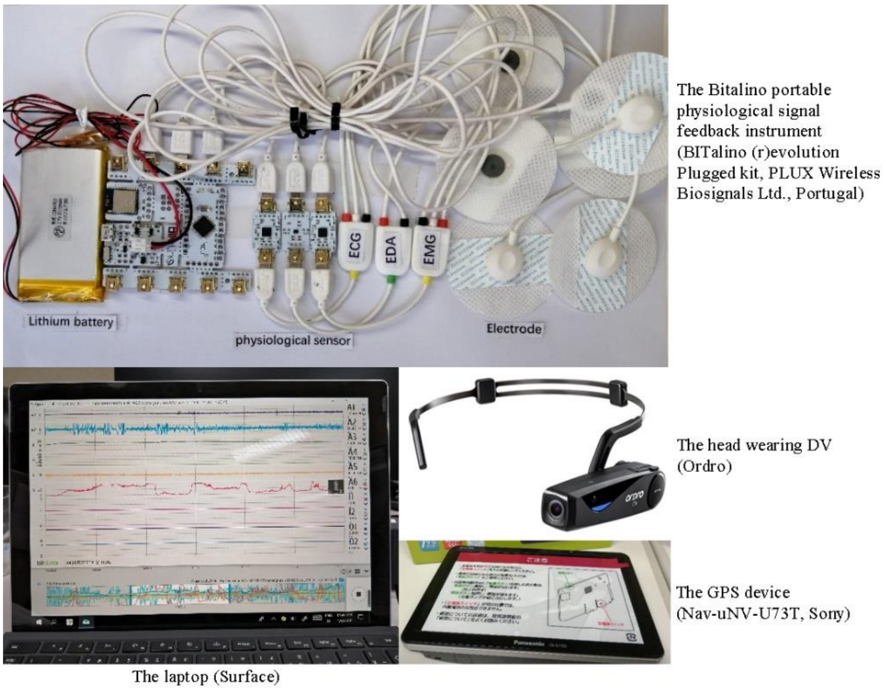


Figure 3.4 Data collection device: a Bitalino portable physiological signal feedback instrument (BITalino (r)evolution Plugged kit, PLUX Wireless Biosignals Ltd., Portugal), a head-wearing DV(Ordro), a GPS device (Nav-uNV-U73T, Sony), and a laptop (Surface).

Figure 3.5 Physiological signals collected in space (A1: EMG signal; A2: ECG signal; A3: EDA signal) stored on a laptop through the supporting software OpenSignals_(r)evolution_win64.

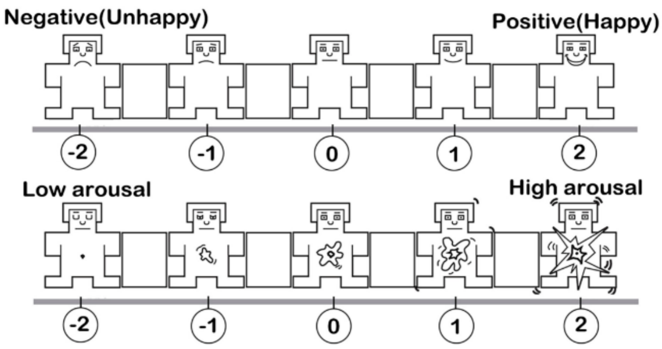


Figure 3.6 Self-Assessment Manikin (Source: Lang, P. J. (1980). Self-assessment manikin. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida).

Table 3.3 Questionnaire on the perception of spatial features.

X_n	**Left adj.**	V_n^{-3}	V_n^{-2}	V_n^{-1}	V_n^{0}	V_n^{+1}	V_n^{+2}	V_n^{+3}	**Right Adj**
		Entirely	Very	Fairly	Neutral	Fairly	Very	Entirely	
X_1	Public	□	□	□	□	□	□	□	Private
X_2	Natural	□	□	□	□	□	□	□	Artificial
X_3	Modern	□	□	□	□	□	□	□	Historical
X_4	Open	□	□	□	□	□	□	□	Enclosure
X_5	Diversity	□	□	□	□	□	□	□	Monotonous
X_6	Easy to identity	□	□	□	□	□	□	□	Uneasy to identity
X_7	Green-rich	□	□	□	□	□	□	□	Insufficient green
X_8	Unique	□	□	□	□	□	□	□	Common
X_9	Beautiful	□	□	□	□	□	□	□	Ugly
X_{10}	Meaningful	□	□	□	□	□	□	□	Meaningless
X_{11}	Artistic	□	□	□	□	□	□	□	Inartistic
X_{12}	Continuous space	□	□	□	□	□	□	□	Interrupted space

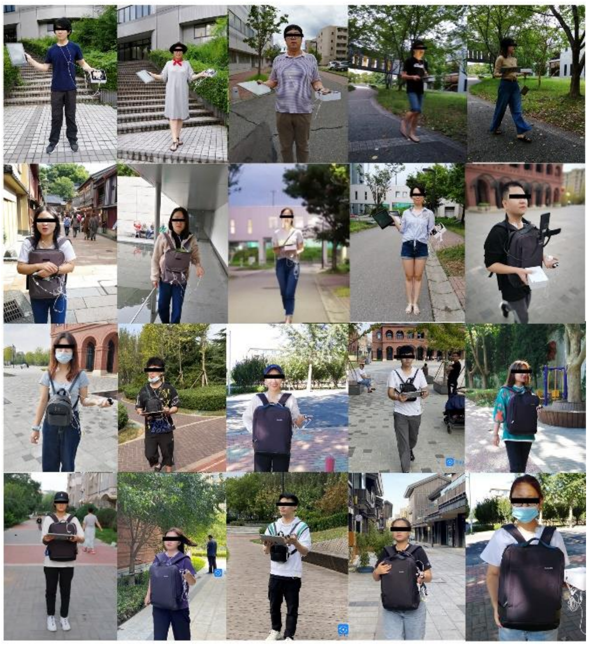


Figure 3.7 Participants walked through the routes of ten sites.

Figure 3.8 The participants wore a physiological signal feedback device (BITalino (r)evolution Plugged kit) and a head-wearing DV (Ordro). They carried a GPS device (Nav-u NV-U73T, Sony) and a laptop computer as they walked through the space routes at a natural pace and constant speed. They could turn their heads to look at their surroundings. The computer automatically recorded the EDA, ECG, and EMG. The position where the electrodes of the physiological signal feedback instrument were pasted on the body: the two EDA electrodes were fixed on the first phalanx of the index and middle fingers; the ECG electrodes were fixed at the carotid arteries on both sides of the neck; the EMG electrodes were fixed on the inner side of the forearm of the arm.

3.3 Data analysis methods

To obtain knowledge from the collected data, we mainly used six methods for data analysis, including Principal component analysis (PCA), Categorical principal components analysis (CATPCA), Entropy weight method (EWM), Ensemble learning, Synthetic minority oversampling technique (SMOTE), Image semantic segmentation, and Fully convolutional neural network algorithm (FCN), Two-step clustering algorithm.

1) Principal component analysis (PCA), Categorical principal components analysis (CATPCA), and Entropy weight method (EWM)

The principal component analysis (PCA) is a commonly used method to reduce the dimensions of continuous variables and perform dimensional analysis. It is usually used to reduce the dimension of a data set by converting a large number of variables into fewer variables that still contain most of the information in the set.

Categorical principal component analysis (CATPCA) applies to data reduction when variables are classified (such as ordinal numbers). The main advantage of using CATPCA instead of traditional PCA is to model the nonlinear relationship between variables, and CATPCA does not require the assumption of multivariate normal data.

The entropy weight method (EWM) is a commonly used weighting method. Its basic idea is to determine the weight according to the size of index variability. It uses the information entropy formula proposed by the information scientist Shannon to calculate each index's entropy weight, then modifies the entropy weight according to each index to obtain an index weight.

Since the results of principal component analysis lack reliable interpretation significance, the entropy weight method can judge the randomness and disorder degree of an indicator by calculating the entropy value. Therefore, we combined the PCA and the entropy weight method to calculate the weights of features, which can eliminate personal factors and reduce the impact of data correlation.

2) Machine learning with single classifier and ensemble learning

Single-classifier machine learning uses one classifier. Ensemble learning combines a variety of classifiers to achieve a better overall classification effect through combining different types of classifiers. The related studies in Table 2.1 mostly use single classifiers (in the orange rectangle), such as Support Vector Machine (SVM), K-Nearest Neighbor algorithm (KNN), Naïve Bayesian (NB), REP-Tree, and Multi-layer Perceptron (MLP), one of them used the ensemble learning classifier Random Forest (RF). The advantage of ensemble learning is that it can improve prediction performance. The three main classes of ensemble learning methods are bagging, stacking, and boosting. To compare the performance of these classifiers, we used three single classifiers and three ensemble classifiers for the model training. The single classifiers were LR, DT, and ANN, and the three ensemble classifiers were DT C5.0 (boosting), RF (bagging), and NN (boosting). They commonly use single and ensemble classifiers and have good recognition performance (Kanjo, Younis and Sherkat, 2018; Kanjo, Younis and Ang, 2019; Ojha et al., 2019; Chen et al., 2021a; Keelawat et al., 2021).

3) Synthetic Minority Oversampling Technique (SMOTE)

SMOTE is a method to address imbalanced datasets by oversampling the minority class. SMOTE works by selecting close samples in the feature space, drawing a line between samples in the feature space, and then drawing new samples at points along the line. First, randomly select a sample from the minority class. Then find the k nearest neighbors of that sample (usually k=5). Neighbors are chosen randomly, and synthetic samples are created at randomly selected points between two samples in the feature space. The researchers found that the method is effective after practical application. A disadvantage of this approach is that synthetic samples are created without considering the majority class, which can make noise affecting the model if the classes overlap strongly.

In this research, we used the SMOTE to solve the problem of imbalanced data collected in public spaces and find adjacent samples and synthesize new minority samples to keep the number of minority samples consistent with the number of majority samples.

4) Image Semantic Segmentation based on Fully Convolutional Neural Network (FCN)

Semantic Segmentation is an important part of image processing and understanding in machine vision technology, and it is also an important branch in the field of AI. Semantic segmentation is to classify each pixel in the image and determine the category of each point (such as background, person or car, etc.) to divide the region. Semantic segmentation has been widely used in scenarios such as automatic driving and UAV landing point determination.

Fully Convolutional Networks (FCN) now use various image semantic segmentation algorithms. FCN replaces the fully connected layer behind the traditional CNN with a convolutional layer. At the same time, to solve the reduction of image size caused by convolution and pooling, the image size is restored by upsampling. The FCN network structure is mainly divided into two parts: the full convolution part and the deconvolution part. The full convolution part is some classic CNN networks (such as VGG, ResNet, etc.), which are used to extract features; the deconvolution part is to obtain the original size of the semantic segmentation image through upsampling. FCN can accept input images of any size and use the deconvolution layer to upsample the feature map of the last convolution layer to restore it to the same size as the input image so that a prediction can be generated for each pixel. At the same time, The spatial information in the original input image is preserved, and finally, pixel-wise classification is performed on the up-sampled feature map. The disadvantage of FCN is that the segmentation results are not refined enough. The image is too blurry or smooth, and the details of the target image are not segmented.

In this research, we used this technology to classify each pixel in the image and determine the category of each point (such as tree、building, sky, car, etc.) to divide the region and then calculate the proportion of each category.

5) Two-step clustering algorithm

The two-step clustering algorithm is unsupervised. It needs to go through two steps to achieve data clustering. The first step is pre-clustering. The samples are roughly divided into several subclasses in a "sequential" manner. The initial stage treats all data as one broad category. Read in - After a sample data, according to the "closeness degree," decide whether the sample should be derived from a new class or merged into an existing subclass. This process will be repeated, eventually forming several classes. The second step is clustering. Based on pre-clustering, subclasses can be merged according to the "degree of closeness" and finally form L' class. It can be seen that this step is a process of decreasing the number of clusters, and with the progress of clustering, the differences within the class will continue to increase. Two-step clustering can simultaneously process numerical and categorical variables; it can determine the number of clusters according to specific criteria; it can diagnose outliers and noise data in samples.

In this research, we used the two-step clustering algorithm to classify the physical and image features of the extracted space.

3.4 Summary

This chapter presents this thesis's research framework and the data collection and analysis methods. We selected ten public spaces in 5 categories in Japan and China as data collection sites. Five types of data, including physiological signals, emotions, space’s physical images, and the user’s perceptual features, are collected through the participants walking in the space. These data were used to build emotional models, feature extraction of high and low-quality public spaces, and comparison of day-to-day spatial features. To achieve these goals, we introduce six data processing and analysis techniques, which help to improve the effect of data mining compared with the related research methods.

Chapter 4
Emotion Classification Models with Physiological Signals

4.1 Introduction

The main goal of this chapter is to build an emotion classification model for public spaces based on physiological signals. The establishment of an emotion classification model based on physiological signals is divided into the following steps: data collection, data processing, feature extraction, model training, and model validation. In related research, after collecting the physiological signal data of participants in a specific route and extracting features, the researchers applied SVM, DT, RF, and other classifiers to establish binary, ternary, and quinary emotional classification models. This model can only adapt to one type of space and is easily affected by specific space elements and features, so the model has limited adaptability. At the same time, these studies did not consider the small sample of participants' negative emotions, resulting in an imbalanced data sample. What affects the model is the specific power. In addition, related studies have not conducted external validation and cannot obtain the final recognition performance of the models. Faced with the above problems, we tried to build the emotion classification models of multi-type public space and improve the classification models' performance using three kinds of physiological signal data, SMOTE sample data balance technology, ensemble classifier, and external verification.

4.2 Methods

Urban public spaces have several types and rich spatial features, and citizens have various demands for these spaces. These two factors make it difficult for designers and managers to make decisions.

There are two methods to evaluate the quality of urban public space: expert evaluation and user evaluation. Because the evaluation indexes of the two methods are different, it is usually not easy to integrate them into a framework. The emotion-eliciting quality evaluation belongs to the second way, which is a quality evaluation method based on the user's emotional experience. This method combines the physical attributes, social attributes, and human perception evaluation of space with the emotional response index.

Related studies in cognitive psychology have argued that external stimuli will impact people's emotions, accompanied by physiological signal changes, which can be regarded as indicators of emotion changes (Kim, Bang and Kim, 2004; Jang et al., 2015; Zhou, 2017). People in urban public spaces are affected by various stimuli: environment, people, cars, commercial activities, etc. Different Spaces stimulate different emotional and physical responses. Conversely, changes in physiological signals caused by physiological responses can be used as a representation of emotional changes.

Researchers in psychology and cognitive computing science use pictures, videos, sounds, and other stimuli to produce physiological responses. And then, they used instruments to collect participants' physiological signal data and establish emotional recognition models through data processing and feature extraction. We applied this research method to the public space; the spatial emotion recognition model can be built and finally used to evaluate the quality of new spatial emotion.

This study used a new approach to evaluating urban public spaces. The main processes conclude obtaining user physiological signals using portable physiological sensors, conducting feature extraction, reduction, and fusion, and using machine classifiers to build emotional recognition models. Finally, proposing a space-positive emotion coupling index system for space design and municipal management decision-making.

However, as the stimulation of public space is usually the weak stimulation of daily life, it will also cause more data noise, resulting in the low recognition rate of the model established by physiological signals. In addition, researchers use different physiological signals as indicators, resulting in a lack of comparability between relevant studies. Some urban design and geography researchers choose a single spatial type to collect data, which limits the generalization performance of the built models (Olsen and Torresen, 2016; Kalimeri and Saitis, 2016; Kanjo, Younis and Sherkat, 2018; Kanjo, Kanjo, Younis and Ang, 2019; Ojha et al., 2019). Therefore, We will use multiple physiological signals to build a effective spatial emotion classification model using ensemble learning through data collection in real space.

User's emotional and spatial feature research are two aspects of this research. This method helps to understand users' spatial experiences and needs from the theoretical level and obtains research results that can be directly applied in urban spatial design and municipal management.

(1) Compared to films, music, drama, and other media, we found that the user's emotional stimulation in urban space is usually weaker, so emotion generated by the former can be called strong emotion, while the latter is weak emotion. Strong emotions are often dotted in urban spatial nodes and opposite locations to form an exemplary sequence of spatial emotional stimulation. Weak emotions depend on the rationality of safety, sanitation, and essential basic functions in urban spaces.

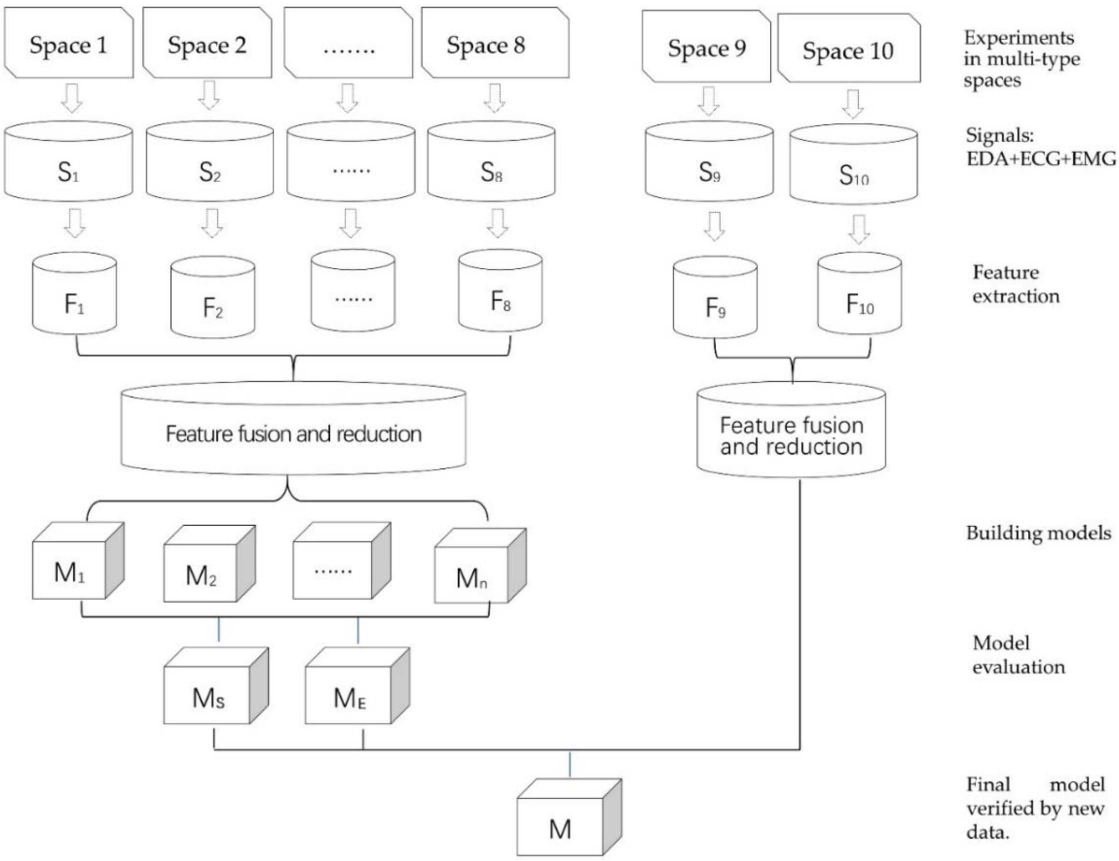
(2) Managers and experts mainly did construction evaluations of urban space in the past. In the future, urban managers can obtain citizens' assessments of urban space quality through smart wristbands, smartphone Apps, and GPS information to ensure that citizens can actively participate in urban environmental management.

(3) The industries of urban design, landscape design, architectural design, and interior design have ever-increasing evaluation requirements at each stage of the project, including environmental evaluation before construction, simulation evaluation in design, and built project evaluation, especially for commercial and tourist projects. Therefore, this professional urban environmental quality evaluation service can become a technical service independent of the design company.

In this study, we proposed a method to evaluate the quality of multi-spatial emotions and verify their effectiveness. This approach can help managers and designers understand users' needs and enable designers to improve their designs. In urban space management, the spatial emotion-eliciting quality evaluation process can be applied to evaluate the spatial emotion-eliciting quality of the space to be reconstructed, providing a verifiable basis for decision-making management.

Typical urban spaces were selected as the stimuli to elicit the participants' physiological and emotional signals. We try to build emotion recognition models using signal processing, feature extraction, and reduction. Figure 4.1 shows a flowchart of the study. In this process, we attempted to improve the method of spatial emotion recognition and applied the proposed model to the public space of another city to further verify its effectiveness.

In practice, the results of this study will contribute to public participation in space evaluation and the improvement of built urban spaces. Meanwhile, for planning urban construction projects, this study will benefit from making spatial decisions in advance and avoiding negative urban spaces.



Note: $S_{1...10}$: signals collected in 10 spaces; $F_{1...10}$: features extracted from signals collected in 10 spaces; $M_{1...n}$: classification models; M_S: models with a single classifier; M_E: models with a ensemble classifier.

Figure 4.1 Flowchart of the data collection and data analysis process

4.2.1 Data pre-processing

The first step is to carry out the participants' emotional valence statistics, as shown in Appendix A. Because it is an odd option for emotion measurement tools, we could get three, five, and seven emotion levels. To build a binary classification model, we deleted the samples whose emotional valence was zero and considered emotions whose valence was -2 and -1 as negative emotions and marked them as “-1”; those whose valence was one and two were positive emotions and marked as “1.” In addition, the statistical results of the SAM scale indicated that, compared to the meaning of emotional valence (positive or negative), emotional arousal was less understood by the participants, who found it difficult to distinguish between emotional arousal and psychological stress.

Furthermore, some participants stated that they would experience psychological stress caused by individual differences as they walked through the public space while wearing instruments. Stress can interfere with emotional arousal. Therefore, we did not use emotional arousal.

The second step is to preprocess physiological signals. Noise reduction was necessary because the physiological signals collected in public urban spaces contained some noise. The interference in the ECG signal primarily results from power frequency interference, electrode contact noise, electromyographic noise, and breathing. Therefore, we used a Butterworth filter to low-pass the ECG signals and applied a zero-phase-shift filter to correct the baseline drift. The denoising of the EDA signal included smoothing, denoising, and filtering using a second-order Butterworth filter with a cut-off frequency of 0.3 Hz. The EMG signal is a waveform diagram of the action potential generated by muscle contraction. Because of the influence of the participants’ walking movements, we applied the Blackman window algorithm to the EMG signal for high- and low-pass filtering (5–50 Hz).

4.2.2 Feature extraction

Based on the GPS positioning, we divided each participant’s EDA, ECG, and EMG signals into four segments; thus, each signal had 400 samples. As 22 samples were incomplete, 378 were valid. To ascertain the number and effectiveness of the features, we applied different software packages to extract features from the EDA, ECG, and EMG signals.

First, we used AcqKnowledge (ver. 4.2) (BIOPAC Systems Inc., 2018) to analyze the EDA signal. The steps to extract the features of EDA are (Figure 4.2):

- Pre-processing: conducting the smoothing processing and low pass (Blackman);
- Segmentation processing;
- Obtaining domain features and nonlinear features;
- Obtaining the frequency domain feature: outputting frequency domain features after Fourier transform.
- Obtaining 16 features, including seven time-domain, four frequency-domain and five nonlinear features.

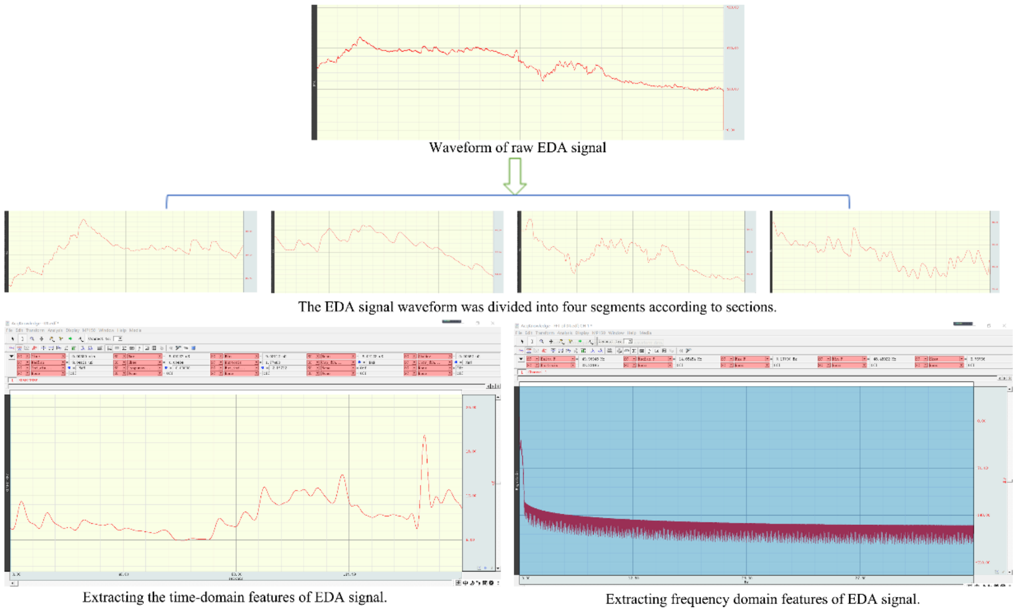


Figure 4.2 Example of feature extraction from EDA signal.

Second is the ECG signal feature extracting. We then used Kubios HRV Premium software (ver. 3.4.3) (Kubios Oy, 2020) to extract the features of the ECG signal. The steps for extracting features are as follows:

- To divide the ECG signal into four sections according to time in the software Acqknowledge 2.4;

- To open the ECG. ACQ file one by one with Kubios software;

- To drag the HR diagram of the second main diagram to the whole diagram display (Note: the length on the left is consistent with the data length on the top);

- To save the features with an Excel format file.

Finally, we obtained 17 time-domain features, 16 frequency-domain features, and 12 nonlinear features (Figure 4.3).

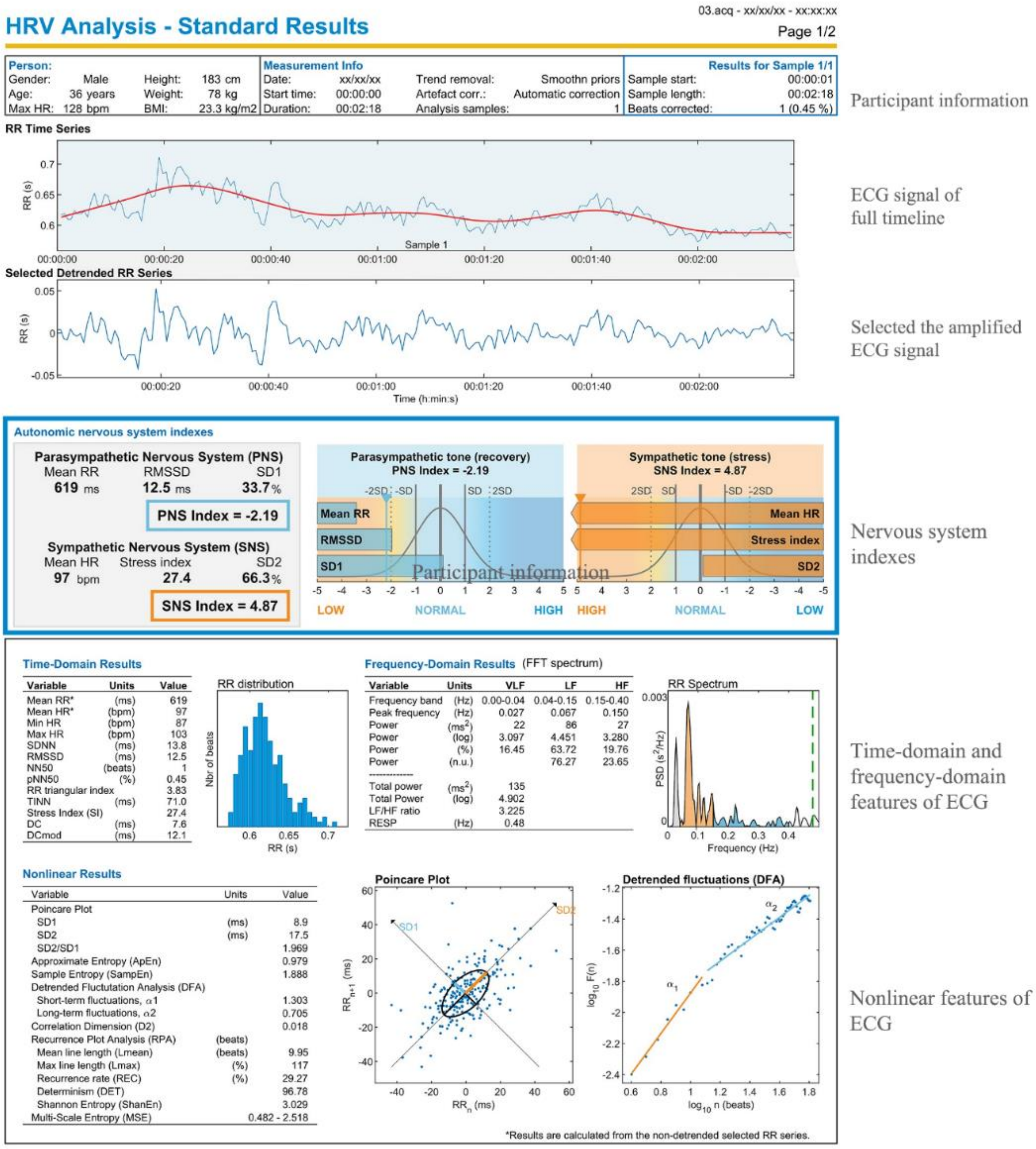


Figure 4.3 Example of feature extraction from ECG signal.

Third, we used the plug-in EMG Toolbar V5.30 (Couturier, 2017) of the software Origin 2019 (OriginLab Corporation, 2019) to extract the features of the EMG signals. The steps to extract the features of EMG are:

- After waveform cutting in Acqknowledge, we save it as a TXT file and store the same signal of each section of each site in the same file as the TXT file;

- Copy the time column and all EMG data columns into the original 2019 (Figure 4.4);

- Select a time column and a signal data column;

- Click the second button of the EMG tool and input the start and end time (Figure 4.5);

- Click the penultimate button (Figure 4.6), and then we obtained five time-domain features and two frequency-domain features.

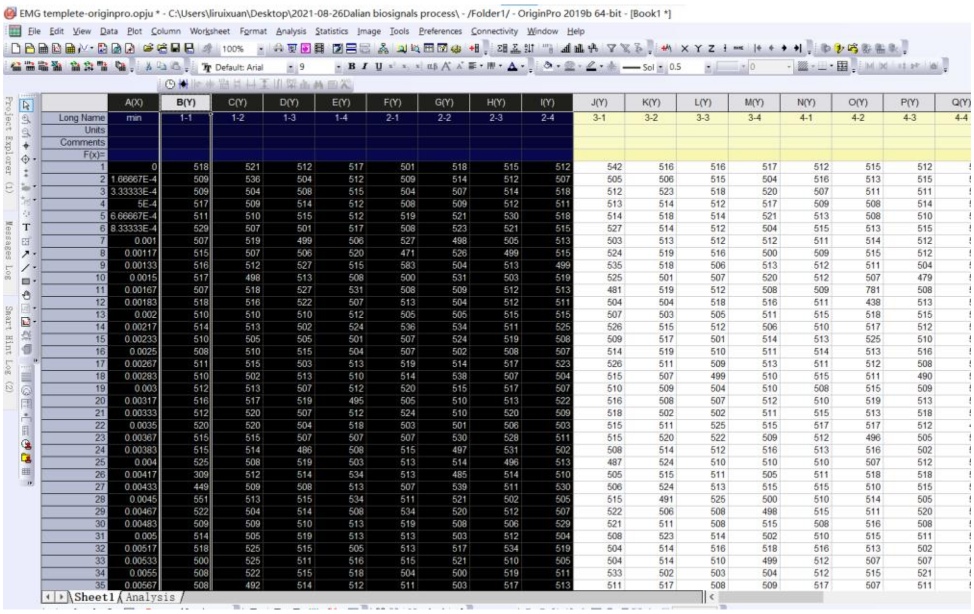


Figure 4.4 EMG data processing using the original 2019.

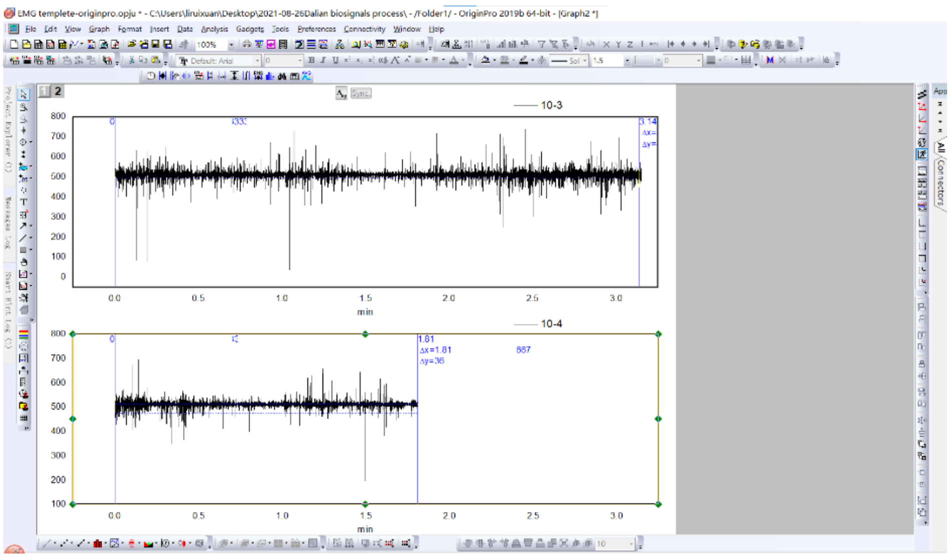


Figure 4.5 Using the EMG tool to process EMG signals.

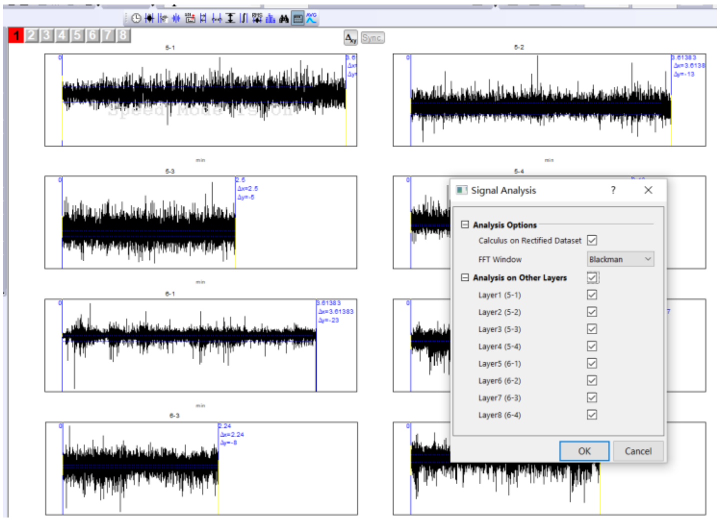


Figure 4.6 Using the EMG tool to extract the features of EMG signals.

Finally, we got 68 signal features after extracting the features of EDA, ECG, and EMG signals.

Feature reduction

We used SPSS (IBM SPSS Statistics 24) to perform PCA on 68 signal features. The results indicated that the significance of the Bartley sphere test was P<0.01, KMO = 0.795, PCA was effective, and the value of extracted eigenvalues was greater than 1 (cumulative% =85.78%) in the components.

After PCA, we will use the entropy weight method (EWM) to calculate the weight of each feature. When each feature's positive and negative directions are not uniform, the data needs to be standardized. The range method can be used for data standardization.

The standardized calculation formula for positive indicators is as follows:

$$X' = \frac{X_{ij} - Min(X_{ij})}{Max(X_{ij}) - Min(X_{ij})}$$

The standardized formula for negative features is as follows:

$$X' = \frac{Max(X_{ij}) - X_{ij}}{Max(X_{ij}) - Min(X_{ij})}$$

Note: X_{ij}: Value of feature i; Max (X_{ij}): maximum value of a feature i; Min (X_{ij}): minimum value of a feature i;

Then we use the entropy weight method's formulas to calculate each feature's weight.

First, to calculate the entropy value E_i of the ith feature i:

$$E_i = -\frac{\sum_{j=1}^{n} p_{ij} \cdot \ln p_{ij}}{\ln n} \quad ; \quad p_{ij} = \frac{X'}{\sum_{j=1}^{n} X'}$$

Then the calculation method of weight w_i is:

$$w_i = \frac{1 - E_i}{\sum_{i=1}^{m} (1 - E_i)}$$

Note: m is the number of features; n is the number of samples.

Then, we had to normalize all the weights to 0~1 because the sum of all the weights is more than 1. Furthermore, if there is negative weight, we need to adjust the positive and negative directions of the factor and re-calculate the wights. Then we use the range method to change each weight to a value between 0-1, and the sum of all weights equals 1. Furthermore, according to the weight, we reduced the number of features. We removed the lower-weight features and selected 50 features (shown in bold text in Table 4.1) highly correlated with emotions.

Table 4.1 Sixty-eight extracted signal features (all) and 50 reduced features (Italics are deleted features).

EDA signal features (8/16)	**ECG signal features (36/45)**		**EMG signal features (6/7)**
Max	Mean RR	HF	Integ.
Min	SDNN	Total power	RMS
Mean	Mean HR	LF/HF ratio	Mean
Stddev	SDHR	SD1	SD
Median	Min HR	SD2	F. mean
Mutual Information	Max HR	SD2/SD1 ratio	F. med.
Median F	RMSSD	Approximate entropy	
Kurtosis	NNxx	Sample entropy	*Max*
	pNNxx	alpha 1	
MinF	RR tri index	Mean line length	
Skew (1)	TINN	Recurrence rate	
Skew	DC	Determinism	
Kurtosis (1)	DCmod	Shannon entropy	
Capacity dimension	AC		
Correlation dimension	ACmod	*VLF (Hz)*	
Information dimension	LF	*LF(Hz)*	
Lyapunov exponent	HF	*HF (Hz)*	
	VLF	*VLF* (ms^2)	
	LF	*VLF (%)*	
	HF	*EDR*	
	LF	*alpha 2*	
	HF	*Correlation dimension*	
	LF	*Max line length*	

4.2.3 Classification model building

We obtained ten datasets, including valence and feature data in ten spaces. We used eight of these for model training and testing. The other two datasets were used as new data to verify the classification capability of the proposed model. We then used SPSS Modeler 18.1 to establish the training and validation models of binary, ternary, and quinary classifications (Figures 4.7, 4.8, and 4.9).

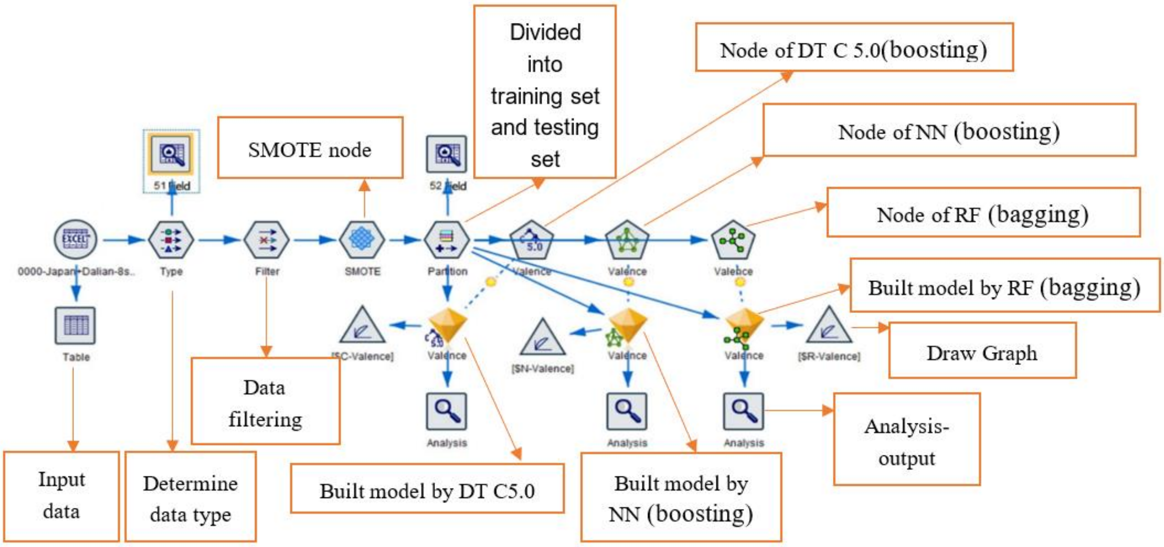


Figure 4.7 Data flow for training and testing the binary classification models.

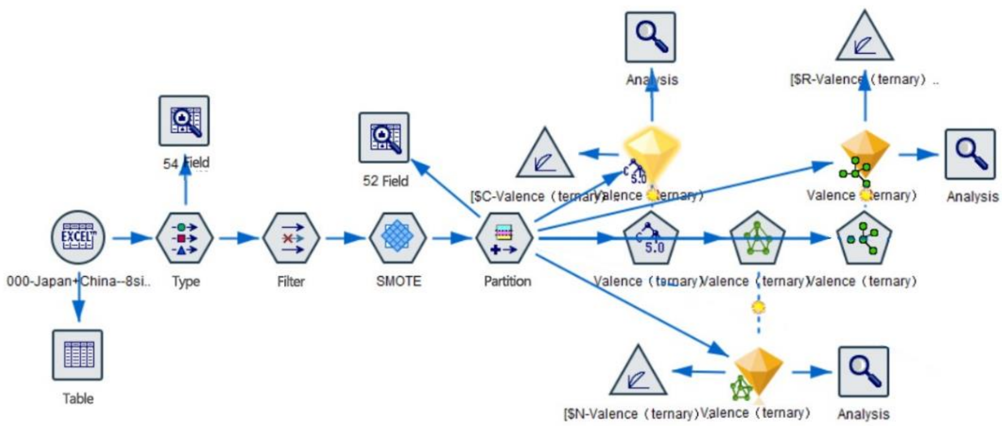


Figure 4.8 Data flow for training and testing the ternary classification models.

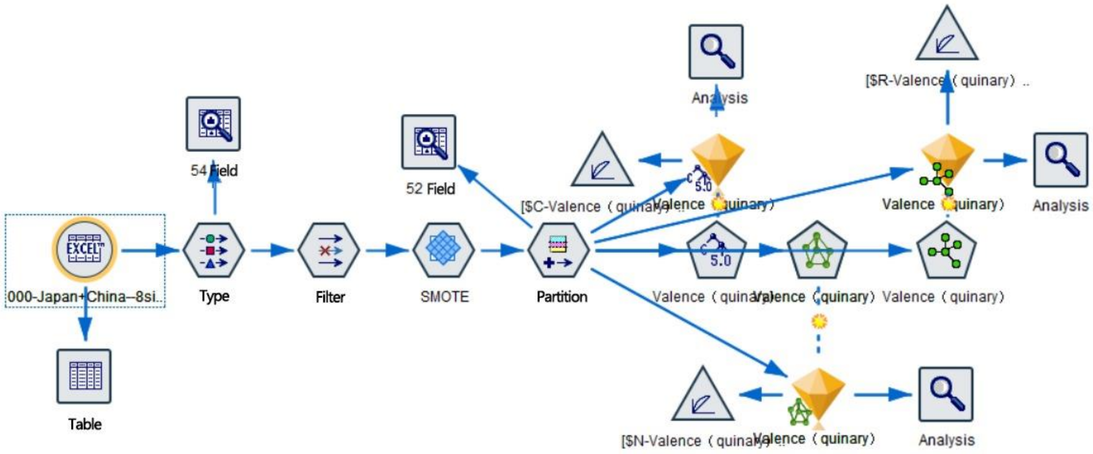


Figure 4.9 Data flow for training and testing the quinary classification models.

Imbalanced data and synthetic minority oversampling technique (SMOTE)

The public space is primarily for citizens' daily leisure and entertainment; thus, the emotions elicited by the space stimulation are mainly positive or calm. Therefore, in the collected data, we observed that the samples of "valence = -2" and "valence = -1" in the dataset were less than others, resulting in poor recognition of negative emotions in the training model. Therefore, we introduced the SMOTE to solve the problem of imbalanced data. Class imbalance refers to an imbalanced distribution of classes in the training set. The proportion of the minority class is equal to or less than 10% of the dataset. When the data is imbalanced, the minority classes do not provide sufficient "information," and the model cannot accurately predict the minority classes. SMOTE is an improved oversampling method (Chawla et al., 2002) that randomly selects an example from a minority group and determines its k-nearest neighbors (KNN) (k = 5 in this example). Subsequently, the algorithm randomly selects a neighborhood in the feature space and a point between the two samples as a new sample, repeats the above steps, and finally balances the majority and minority samples (Figure 4.10).

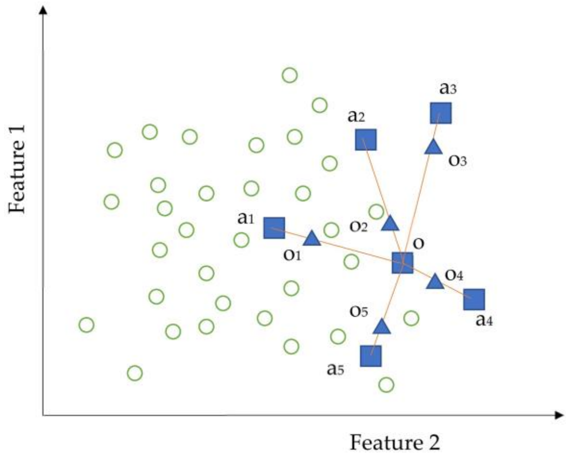


Figure 4.10. SMOTE algorithm: The blue square and green circle represent the minority and majority classes, respectively. The KNN of point O in the minority set was obtained by calculating the Euclidean distance between O and each sample in the set. Based on the k (k = 5), the algorithm connected the k (k = 5) minority points (a1, a2, a3, a4, a5) around O and finally inserted new synthetic points (O1, O2, O3, O4, O5) on the line of the two points until the number of all the minority types and insertion points was balanced with the number of majority types.

Single and ensemble classifiers

In related research, single classifiers were used, including LR, SVM, DT 5.0, ANN, and RF ensemble classifiers (Iliou and Anagnostopoulos, 2009; Molavi, Yunus and Akbari, 2012; Kanjo, Younis and Sherkat, 2018; Kanjo, Younis and Ang, 2019; Ojha et al., 2019; Chen et al., 2021a; Keelawat et al., 2021). So, related research mainly used single classifiers and one ensemble classifier.

Ensemble learning achieves better predictive performance by combining predictions from multiple models. The three main classes of ensemble learning methods are bagging, stacking, and boosting. Among these, bagging and boosting are used more often than stacking. Bootstrap aggregation (bagging) is an ensemble learning method that achieves a diverse group of ensemble members by varying the training data. Boosting is a machine learning algorithm that can be used to reduce deviations in supervised learning. Boosting learns a series of weak classifiers and combines them into a robust classifier.

To improve the model's ability, we selected ensemble classifiers, and to compare the ability between single classifiers (used by related researchers) and ensemble classifiers, we used three single classifiers and three ensemble classifiers for the model training. The single classifiers were LR, DT, and ANN, and the three ensemble classifiers were DT C5.0 (boosting), RF (bagging), and the neural network (boosting).

4.2.4 Model evaluation

Selection of evaluating indicators of the models

The confusion matrix, also known as the error matrix, is a standard format for accuracy evaluation. It can be used to calculate the performance indices of the classification model: accuracy, recall, and F1-score. The calculation method for each index is as follows.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

$$Recall(R) = TP/(TP + FN)$$

$$F1 - Score = 2P * R/(P + R) \qquad P(precision) = TP/(TP + FP)$$

Note:

TP = No. of true positives among total predictions; FP = No. of false positives among total predictions;

FN = No. of false negatives among total predictions; TN = No. of true negatives among total predictions.

In addition to the above three indices, we also selected the area under the curve (AUC) and the Gini coefficient as the performance indices of the binary classification model. The AUC is a popular measure of the degree or measure of separability. This indicates the extent to which the model can distinguish between the two classes. The value range of the AUC is between 0.5 and 1. An AUC of 0.5 indicates the worst performance. The closer the AUC is to 1.0, the better the model's performance.

$$AUC = \frac{\sum_{i \in positiveClass} rank_i - \frac{M(1+M)}{2}}{M \times N}$$

Note: M is the number of positive samples, N is the number of negative samples, and $rank_i$ is the serial number of sample i.

The Gini coefficient compares the Lorenz curve of a ranked empirical distribution to the line of perfect equality. It measures the degree of concentration (inequality) of a variable within the distribution of its elements. It is calculated as follows:

Gini coefficient = area A / (area A + area B) = twice the area A

Note: Area A is the area between the line of equality and the Lorenz curve. Area B is the remaining area between the Lorenz curve and the x axis.

For the indices of the ternary and quinary class classification models, we also selected Cohen's kappa coefficient to test the consistency of the classification results. Cohen's kappa is a statistical coefficient representing the classification's degree of accuracy and reliability. It measures the agreement between two raters who classify items into mutually exclusive categories (Uebersax, 1982). The kappa value is always less than or equal to one, indicating less-than-perfect or perfect agreement, respectively. Cohen's kappa coefficient was calculated as follows:

$$k = (p_o - p_e)/(1 - p_e)$$

Note: where p_o is the relative observed agreement among raters, and p_e is the hypothetical probability of chance agreement.

External validation

In addition to internal testing, the performance of the models was subjected to external validation. We input the two previously selected spatial datasets (collected from Japan and China) into the built binary, ternary, and quinary classification models to verify the model's effectiveness at predicting new spatial emotion-eliciting quality (Figure 4.11). The models output results for the two spaces. By comparing the output classification results with the raw valence values, we obtained the accuracy and confusion matrices of the classification.

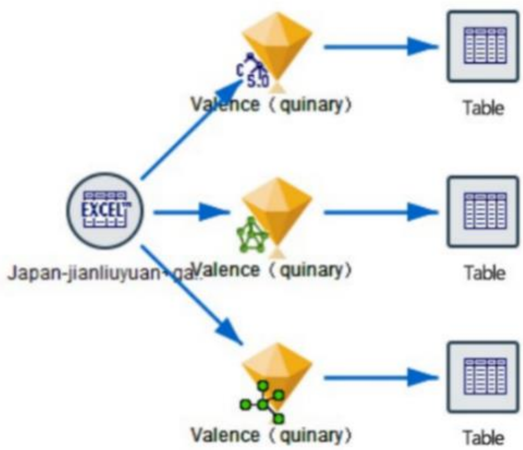


Figure 4.11 Data flow for external validation.

4.3 Results

4.3.1 The effect of feature reduction on the models

The PCA algorithm was used to reduce the extracted 68 features to 50. However, although the PCA algorithm reduced the dimension of the independent variables, the significance of these independent variables to the target variable was unclear. To verify whether the reduction in the number of features had a positive effect on valence classification, we used 68 and 50 signal features to build binary and ternary classification models, respectively (RF (bagging) and ANN (boosting) as classifiers). Table 4.2 presents the model performance results before and after feature reduction.

Table 4.2 Comparison of the performance of the models based on 68 features and 50 features.

Class	Classifier	68 features					50 features				
		Accuracy	Recall	Fl-Score	AUC	Kappa	Accuracy	Recall	Fl-Score	AUC	Kappa
Binary	RF (bagging)	75.90%	0.780	0.755	0.886		83.30%	0.833	0.816	0.951	
	ANN (boosting)	85.20%	0.855	0.855	0.879		92.60%	0.929	0.929	0.962	
Ternary	RF (bagging)	87.10%	0.870	0.870		0.804	91.10%	0.917	0.910		0.866
	ANN (boosting)	83.60%	0.823	0.827		0.750	90.20%	0.907	0.900		0.852

4.3.2 Classification results and performance comparison

Binary classification

We divided the eight datasets used for the training and testing models into two parts, in the ratio of 8:2, randomly selected as the training and test sets, respectively. The values of the target variable for binary classification were "-1, 1," and 50 signal features as the independent variables. The model evaluation results are presented in Table 4.3.

Table 4.3 Evaluation result comparison of binary classification models with six classifiers.

Classifers	Recall	Fl-Score	AUC	Gini	Accuracy
LR	0.679	0.679	0.718	0.436	74.29%
DT C5.0	0.571	0.571	0.642	0.285	65.71%
ANN	0.964	0.931	0.917	0.833	94.29%
DT C5.0 (boosting)	0.750	0.706	0.888	0.777	72.22%
RF (bagging)	0.833	0.816	0.932	0.865	83.33%
ANN (boosting)	0.929	0.929	0.971	0.942	92.59%

The results of binary classification indicated that the recognition accuracies of the models based on the ANN and ANN (boosting) were higher than 90%, and they had better classification performance. These results also indicate that the two models effectively evaluated the emotion-eliciting quality evaluation of urban public spaces.

Ternary classification

The value of the target variable for ternary classification were “-1, 0, and 1”, and all the valid sample data were used in model training or testing. The sample data were divided into training and test sets at a ratio of 8:2, and SMOTE was used for data over-sampling. After testing the models, we obtained the classification accuracy and average of each class of model performance index, as presented in Table 4.4 and Figure 4.12.

Table 4.4 Performance comparison of ternary classification models with six classifiers.

Classifiers	Recall	Fl-Score	Kappa	Accuracy
LR	0.620	0.623	0.432	62.20%
DT C5.0	0.890	0.877	0.808	87.40%
ANN	0.770	0.770	0.642	76.38%
DT C5.0 (boosting)	0.767	0.763	0.651	79.17%
RF (bagging)	0.917	0.910	0.866	91.07%
ANN (boosting)	0.907	0.900	0.852	90.18%

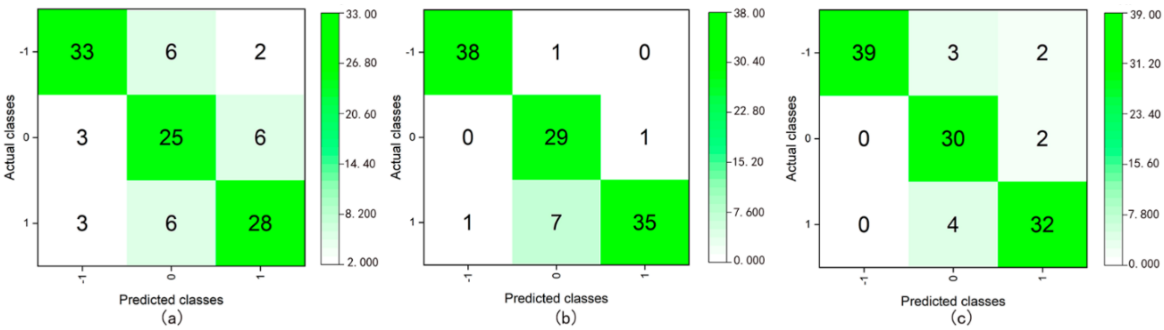


Figure 4.12 Confusion matrices of the ternary class classification using DT C5.0 (boosting) (a), RF (bagging) (b), and ANN (boosting) (c).

The performance indices of each class classification in the ternary classification model are listed in Tables 4.5 and 4.6.

Table 4.5 Performance indexes of each class classification in the ternary classification model with three single classifiers.

Valence	LR			DT C5.0			ANN		
	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy
-1	0.620	0.680	80.31%	0.970	0.880	0.940	0.870	0.900	94.49%
0	0.590	0.560	71.65%	0.870	0.890	0.920	0.780	0.670	80.31%
1	0.650	0.630	72.44%	0.830	0.860	0.890	0.680	0.740	77.95%

Table 4.6 Performance indexes of each class classification in the ternary classification model with ensemble classifiers.

Valence	DT C5.0 (boosting)			RF (bagging)			ANN (boosting)		
	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy
-1	0.800	0.820	87.50%	0.970	0.970	9&21%	0.890	0.940	95.54%
0	0.740	0.700	81.25%	0.970	0.870	91.96%	0.940	0.870	91.96%
1	0.760	0.770	84.82%	0.810	0.890	91.96%	0.890	0.890	92.86%

From the results of the ternary classification, we observed that the models based on the ANN (boosting) and RF (bagging) had higher performance index values, and their recognition accuracies were 91.07% and 90.18%, respectively. Moreover, the models exhibited better classification abilities for each class (Figure 4.12). The results indicated that both models could also effectively evaluate the affective quality of urban public spaces.

Quinary classification

The value of the target variable for quinary classification was "-2, -1, 0, 1, 2", and all the valid sample data were used to build the models. We divided the sample data into training and test sets according to a ratio of 8:2 and used SMOTE for data oversampling. After testing the models, we obtained the classification accuracy and average of Recall, F1-score, and Kappa for each class, presented in Table 4.7 and Figure 4.13.

Table 4.7 Performance comparison of quinary classification models with six classifiers.

Classifiers	Recall	Fl-Score	Kappa	Accuracy
LR	0.562	0.552	0.531	57.43%
DT C5.0	0.594	0.576	0.561	60.32%
ANN	0.624	0.636	0.587	61.12%
DT C5.0 (boosting)	0.696	0.696	0.624	69.86%
RF (bagging)	0.656	0.658	0.584	66.67%
ANN (boosting)	0.582	0.59	0.529	62.41%

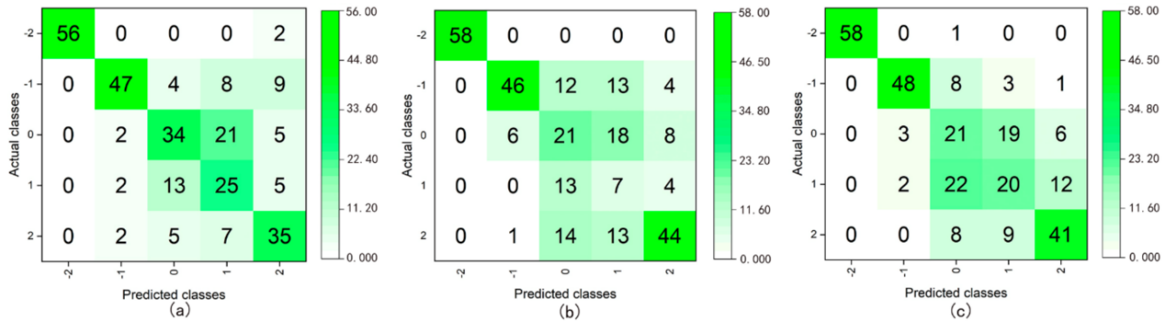


Figure 4.13 Confusion matrices of the quinary class classification using DT C5.0 (boosting) (a), RF (bagging) (b), and ANN (boosting) (c).

The performance indices of each class classification in the quinary classification model are listed in Tables 4.8 and 4.9.

Table 4.8 Performance indexes of each class classification in the quinary classification model with three single classifiers.

Valence	LR			DT C5.0			ANN		
	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy
-2	0.510	0.550	56.61%	0.590	0.580	61.26%	0.520	0.560	60.16%
-1	0.770	0.770	78.20%	0.730	0.710	74.25%	0.740	0.750	76.35%
0	0.450	0.430	45.12%	0.460	0.500	55.77%	0.510	0.480	55.61%
1	0.510	0.500	50.12%	0.510	0.560	56.06%	0.550	0.540	57.56%
2	0.590	0.590	60.01%	0.610	0.690	70.56%	0.540	0.600	62.91%

Table 4.9 Performance indexes of each class classification in the quinary classification model with ensemble classifiers.

Valence	DT C5.0 (boosting)			RF (bagging)			ANN (boosting)		
	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy	Recall	Fl-Score	Accuracy
-2	0.970	0.980	99.29%	0.980	0.990	99.65%	1.000	1.000	100.00%
-1	0.690	0.780	90.43%	0.800	0.850	93.97%	0.610	0.720	87.23%
0	0.550	0.580	82.27%	0.430	0.390	76.24%	0.400	0.370	74.82%
1	0.560	0.470	80.14%	0.360	0.370	76.24%	0.290	0.190	78.37%
2	0.710	0.670	87.59%	0.710	0.690	87.23%	0.610	0.670	84.40%

The results of the quinary classification indicated that the model that incorporated DT C5.0 (boosting) had the best classification performance. However, its accuracy was only 69.86%, and the kappa coefficient was low, which demonstrated that the recognition performance of each class was very uneven, although some classes had 100% accuracy. Thus, in practice, these six models cannot satisfy the quinary classification of the affective quality of a space.

The comparison of the four indices of the binary, ternary, and quinary classification models with the best performance is shown in Figure 4.14. The results indicated that the classification ability declined sequentially, and the quinary class classification ability declined. The binary and ternary class classification models might satisfy the practical requirements.

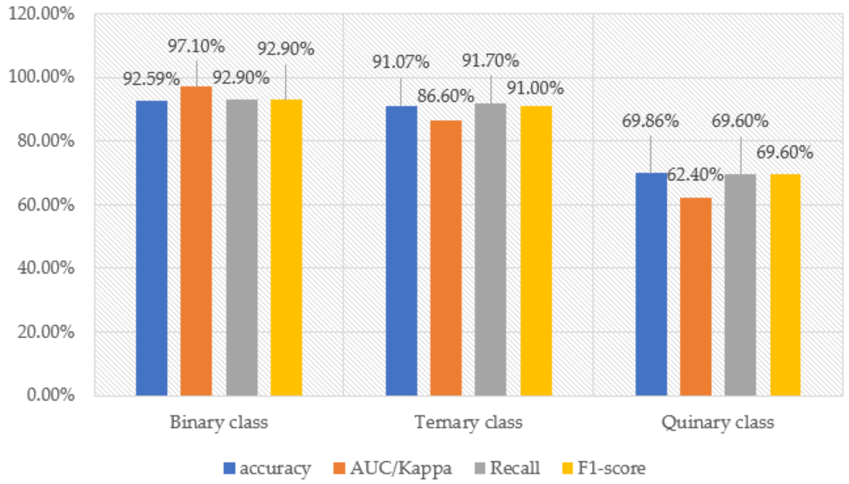


Figure 4.14 Comparison of accuracy and main performance indexes (binary: AUC; ternary and quinary: Kappa) among the binary, ternary, and quinary classification models.

External validation

In addition to internal testing, the performance of the models was subjected to external validation. We input the two previously selected spatial datasets (collected from Japan and China) into the built binary, ternary, and quinary classification models to verify the model's effectiveness at predicting new spatial emotion-eliciting quality. The models output results for the two spaces (Appendix B). By comparing the output classification results with the raw valence values, we obtained the accuracy and confusion matrices of the classification, as shown in Table 4.10, Figures 4.15, and 4.16.

Table 4.10 Classification accuracy of external validation of the two new spaces using the proposed binary, ternary, and quinary-class classification models.

Models	Binary	Ternary	Quinary
LR	46.80%	44.40%	43.10%
DT C5.0	65.90%	53.30%	51.40%
ANN	61.70%	56.90%	55.60%
DT C5.0 (boosting)	65.90%	54.20%	47.20%
RF (Bagging)	78.70%	65.30%	55.60%
ANN (boosting)	80.90%	62.50%	61.10%

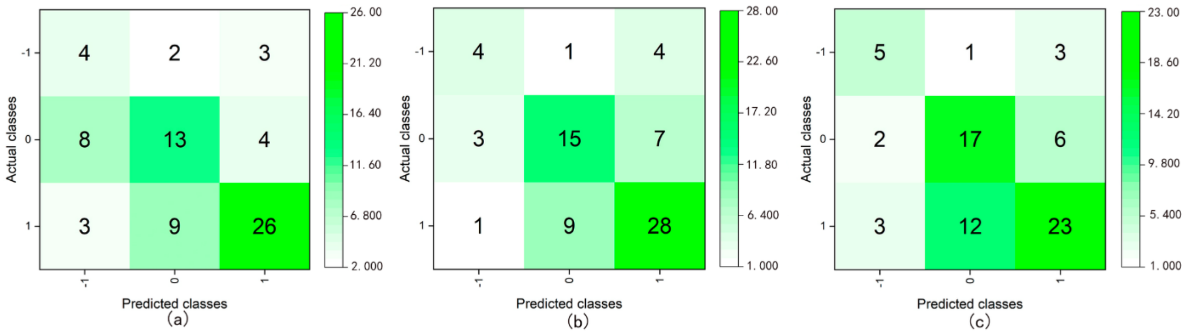


Figure 4.15 Confusion matrices of the ternary class classification for external validation using DT C5.0 (boosting) (a), RF (bagging) (b), and neural network (boosting) (c).

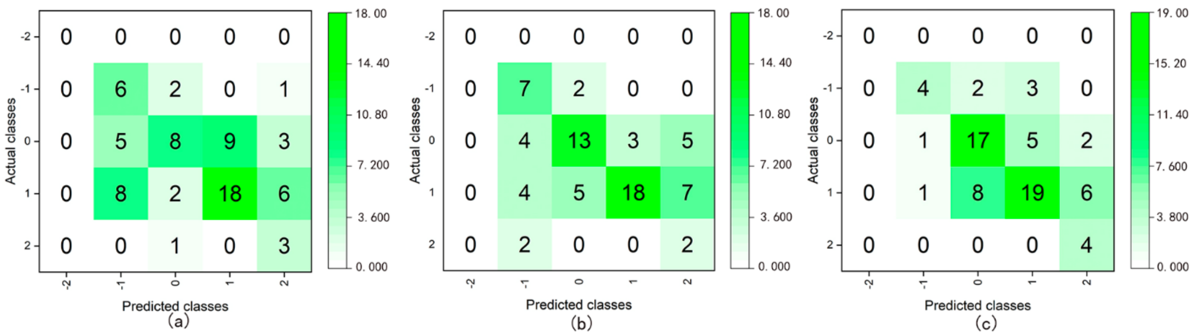


Figure 4.16 Confusion matrices of the quinary class classification for external validation using the DT C5.0 (boosting) (a), RF (bagging) (b), and ANN (boosting) (c).

The results indicated that the highest accuracy of external validation in binary classification was 80.9%, whereas those of ternary and quinary types were 65.3% and 61.1%, respectively.

Moreover, the accuracies of the ensemble classifiers were generally higher than those of the corresponding single classifiers. The confusion matrix of the ternary classification indicated that the classification results of samples whose valences were “-1” were lower than those of the other classes. Because there was no sample whose valence was “-2” in the new data, the quinary classification result was zero, and the classification results of the samples whose valences were zero and one were more accurate than those of the others.

4.4 Discussion

This study built and examined models suitable for evaluating the affective quality of multitype public spaces. We improved the model’s performance through feature selection, SMOTE, and ensemble classifiers and used external validation to verify the model's actual performance.

First, as shown in Table 2.1, previous researchers extracted 8–188 features from physiological signals. This considerable difference in the number of features was owing to the difference in the number of physiological signals and the feature extraction method. Therefore, to ensure the comparability of the studies and facilitate their operation in practical applications, we selected three commonly used physiological signals, EDA, ECG, and EMG, and the PAC method, which is widely used to reduce the feature dimensions. As shown in Table 4.2, with the same classifier, the recognition accuracy of the model increased by 6.35% on average after the number of features was reduced from 68 to 50; other indices improved as well. These results indicated that the PAC algorithm effectively eliminated data redundancy and noise and improved the classification ability of the model. However, obtaining a definite number of features remains a challenge and solving this problem requires scholarly consensus following extensive experiments.

Second, Compared with positive emotions, fewer spaces elicit negative emotions unless they are undeveloped or under problematic management. Thus, we had a situation where the data sample contained insufficient examples of negative emotions, occasionally less than 1/10 of the positive emotion samples. The imbalanced samples resulted in inaccurate predictions. Generally, up-sampling and down-sampling the data or algorithm level can solve this problem; however, simply increasing the amount of data by duplication affects a model’s adaptability. On the other hand, directly reducing the sample size results in information loss. Oversampling techniques, such as SMOTE, increase the number of minority samples. Additionally, it has minimal effect on the information in the data, making it possible to obtain a model with better classification ability.

Third, we used three ensemble classifiers and compared their performances with those of single classifiers. In the past ten years, ensemble classifiers have demonstrated strong classification performance. Compared with the models established using classifiers, such as SVM (Olsen and Torresen, 2016; Kanjo, Younis and Sherkat, 2018; Ojha et al., 2019), KNN (Kanjo, Younis and Sherkat, 2018), BEP-tree (Ojha et al., 2019), MLP (Ojha et al., 2019), and RF (Kalimeri and Saitis, 2016; Kanjo, Younis and Sherkat, 2018) in related studies, the ensemble classifiers used in this study exhibited a higher classification accuracy. We observed that the average accuracy of the ensemble classifiers was 7.59% higher than that of the single classifiers. Comparing the Gini and Kappa coefficients yielded similar results, indicating that these ensemble classifiers adapted better to the multi-noise data collected in urban public spaces.

Table 4.11 compares the average accuracies of binary, ternary, and quinary classification models with single and ensemble classifiers, and the table shows that the average accuracies of the ensemble classifiers are higher than that of the single classifiers.

Table 4.11 Comparison of the average accuracies of binary, ternary, and quinary classification models with single and ensemble classifiers.

	Classifiers	Recall	Fl-Score	Gini	Accuracy	Average of accuracy
Binary classification	LR	0.679	0.679	0.436	74.29%	78.09%
	DT C5.0	0.571	0.571	0.285	65.71%	
	ANN	0.964	0.931	0.833	94.29%	
	DT C5.0 (boosting)	0.75	0.706	0.777	72.22%	82.73%
	RF (bagging)	0.833	0.816	0.865	83.33%	
	ANN (boosting)	0.929	0.929	0.942	92.59%	
Ternary classification	LR	0.62	0.623	0.432	62.20%	75.33%
	DT C5.0	0.89	0.877	0.808	87.40%	
	ANN	0.77	0.77	0.642	76.38%	
	DT C5.0 (boosting)	0.767	0.763	0.651	79.17%	86.81%
	RF (bagging)	0.917	0.91	0.866	91.07%	
	ANN (boosting)	0.907	0.9	0.852	90.18%	
Quinary classification	LR	0.562	0.552	0.531	57.43%	59.62%
	DT C5.0	0.594	0.576	0.561	60.32%	
	ANN	0.624	0.636	0.587	61.12%	
	DT C5.0 (boosting)	0.696	0.696	0.624	69.86%	66.31%
	RF (bagging)	0.656	0.658	0.584	66.67%	
	ANN (boosting)	0.582	0.59	0.529	62.41%	

Fourth, external validation is a method for validating the predictive ability of a model by entering a new dataset. Related studies have shown that good test results do not guarantee that a model will have good adaptability. The model's predictive ability for new data is often lower than the test results (Consonni, Ballabio and Todeschini, 1982; Vergouwe et al., 2005; Collins et al., 2014). Similar results were obtained in our study.

The results of the external validation of the quinary classification were worse than those of the test results. We attributed this to using different spatial data and participants and the limited sample size of external verification. Quinary classification requires a larger sample size than binary and ternary classification.

Fifth, a comparison of Figures 4.13 and 4.16 reveals that the two classification results were almost the opposite. In the classification of the test set, the classification results of the samples whose valences were “-2”, “-1”, and “2” were better than that of others. In contrast, the classification results of the samples whose valences were “0” and “1” were better than that of others in the external validation classification. This may be owing to the use of SMOTE, which increases the minority class samples through oversampling, increases the number of samples with similar information to the original samples, and reduces the model’s ability to classify new minority class samples. The impact of SMOTE was limited in binary and the ternary classifications because of the large sample size. Therefore, external validation was a further step toward verifying the model’s actual performance. Although SMOTE is suitable for large sample sizes, as the number of classes increases, the sample size of each class decreases, and its effect become minimal.

Sixth, the results of related studies show that the highest recognition accuracy of binary classification of the test set is 87% (Table 2-1), the object of that research is a public space, and the data of the training set and testing set are from the same space. The objects of our research are five types of public spaces, so the difficulty in learning and recognition increases due to the differences in spatial properties and functions. Furthermore, the highest recognition accuracy of binary classification of the testing set of our research is 94.29%. In addition to the testing set verification, we also input new data from two other spaces into the built models for external verification, and the highest accuracy of the binary classification is 80.90% (Figure 4.17). No external validation was conducted in related research, so there were no comparable research results.

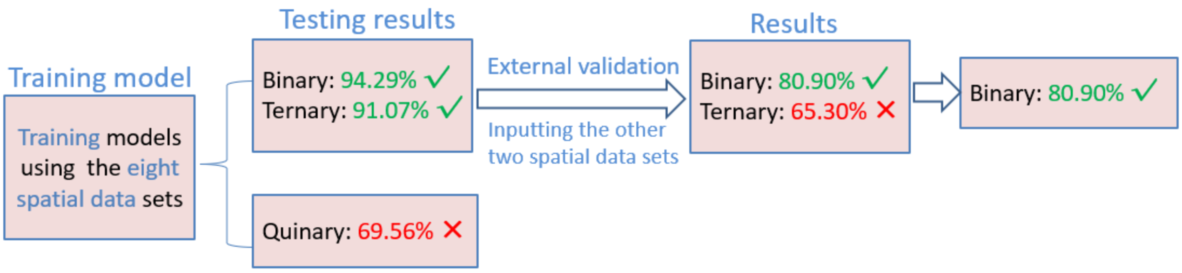


Figure 4.17 The highest accuracies of model testing and external validation.

In addition, since we collected data in real-world space, participants' walking, breathing, surrounding people, cars, etc., would impact the physiological signals of participants; the accuracy of model recognition would be affected. Therefore, the accuracy of external validation of the model declined to 80.90%.

4.5 Summary

This chapter describes the study on building a multi-type spatial binary, ternary and quinary emotion classification model with physiological signal data and ensemble classifiers. Through model evaluation and external verification, the results show that the recognition performance of binary models might meet the actual needs, and the ternary and quinary classification model could not meet the real needs.

Whether through expert or user evaluation, evaluating public spaces in different regions, styles, and functions has always been a controversial problem in urban science. Our focus was on enhancing the classification capabilities of the proposed model. We collected data from five types of spaces in two countries to ensure the diversity of spatial data. In addition, we improved the classification performance of the model using efficient feature reduction, SMOTE algorithm, and ensemble learning. We also compared the performances of the binary, ternary, and quinary classification models. Finally, through external validation, we observed that the binary and ternary classification models outperformed the quinary model in satisfying decision-making requirements on urban public space renewal.

Chapter 5
Association between Spatial Features and Emotions

5.1 Introduction

The main goal of this chapter is to extract quantitative physical and image features of multiple types of positive spaces. The quantitative result of the spatial features that affect users' emotions is the demand for refined urban design and municipal management. In previous studies, some researchers directly proposed the spatial features that affect user experience based on professional backgrounds. In contrast, others asked users which elements or features are attractive and which are not. We tried to extract space physics and image features oriented to practice. Because emotional valence is users' emotional response after a space experience, it can be used as the basis to judge the popularity of space. Then, by looking for the relationship between these two features and the user's emotional valence, we can find the corresponding spatial features of high and low emotional valence as the basis for judging the spatial quality.

5.2 Methods

5.2.1 Spatial features extraction

In terms of spatial feature extraction, previous studies mainly extracted physical space features, commercial activity features, and primary features through questionnaires (Mehta and Bosson, 2018; Rahman et al., 2020). This method lacks consideration of the weight of different features and does not distinguish and compare different spatial features. This study will use Principal Component Analysis (PCA) and cluster analysis to find the corresponding relationships between the weights, spatial features, and users' emotions, and the results will help designers choose suitable spatial features according to the features of the space and provide quantifiable evidence for urban space design.

Physical features

Spatial features usually include space physical, aesthetic, and functional features. Because spatial attributes often determine the function features of space, the results are not generalizable. The aesthetic features of space are mainly obtained through a questionnaire survey and are people's psychological reactions to the physical environment, and the physical features are the carrier of aesthetic features. Therefore, this chapter takes the physical features of space as the primary research object.

Physical features can be divided into two types: indicators that can be directly quantified and difficult to quantify. Directly quantifiable indicators such as the length and width of urban space, the height of buildings, and the height of trees on both sides; The indexes that are difficult to quantify directly include closeness, continuity of space boundary, and permeability of space boundary. This study proposes 13 indicators in six categories by analyzing the results of previous studies (Figure 5.1, 5.2, 5.3, Appendix C and D). The value of each indicator is determined through field measurement and Google earth measurement tools. Although each section has the same attributes, the spatial features of different parts in the same section are not identical. Therefore, we made a section at each section's front, middle and rear sections to determine the value of the features in each section and then calculate the average value. Thus, the physical features of the 40 sections were obtained (Appendixes C and D).

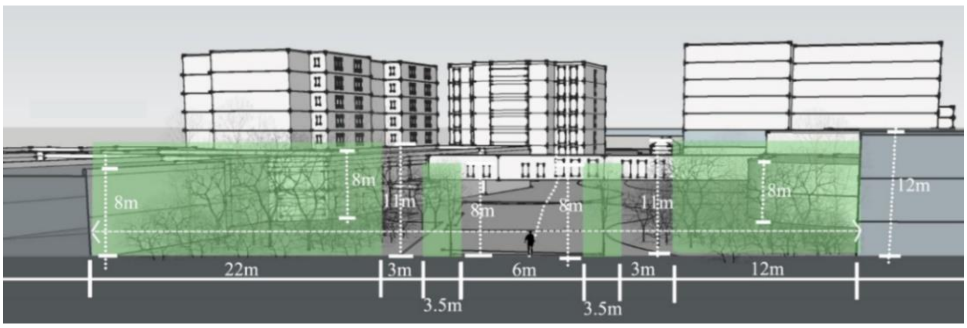


Figure 5.1 Extraction of physical features from one of the sections in the JAIST campus.

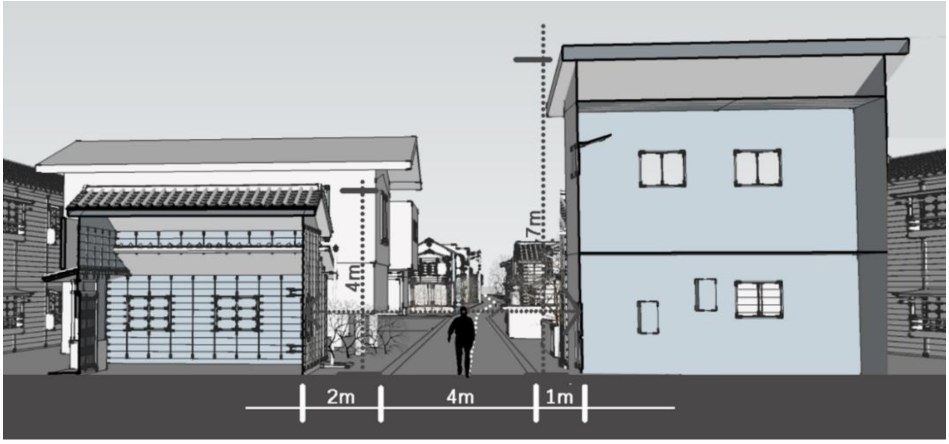


Figure 5.2 Extraction of physical features from one of the sections in the residential area, Yokaichi.

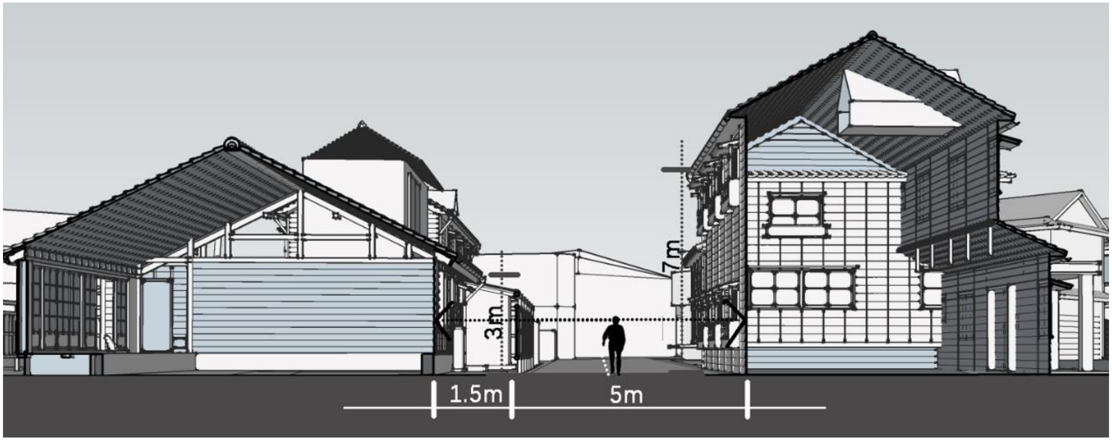


Figure 5.3 Extraction of physical features from one of the sections in the Higashi Chaya District.

Image features

We used GPU-UDA-enabled Semantic Segmentation App. V1.0 (https://github.com/shelhamer/fcn.berkeleyvision.org), a visual image Semantic Segmentation software based on deep learning Full Convolutional Network (FCN) developed by the School of Information Engineering of China University of Geosciences (Wuhan) to conduct automatic image segmentation (Guan, 2019). This software can distinguish the area proportion of 150 objects in outdoor and indoor space images (Appendix E). The training and test set's recognition accuracy are 0.814426 and 0.66839, respectively. Among these 150 objects, we only used it to extract 24 objects related to outdoor public space. After analyzing and merging 24 objects, we finally got 13 outdoor things, as shown below; this choice is similar to the number of subjects in previous studies (Huang et al., 2014; Lin et al., 2021; Zhang et al., 2018; Tang and

Long, 2019; Liu et al., 2017; Yao et al., 2019).

Classification of objects in images (13 classes):

a. Id_2-building and wall = Id_1-wall + Id_2-building+ Id_26-house
b. Id_3-sky
c. Id_5-tree and plant = Id_5-tree + Id_18-plant
d. Id_7-road and sidewalk = Id_7-road + Id_12-sidewalk + Id_14-earth + Id_53-path+ Id_4-floor
e. Id_10-grass
f. Id_13-person
g. Id_21- car
h. Id_22-water = Id_22-water + Id_61-river + Id_129-lake
i. Id_33-fence = Id_33-fence + Id_96-bannister
j. Id_35-rock
k. Id_44-signboard
l. Id_88-streetlight
m. Id_94-pole = Id_94-pole + Id_43-column

The following Figure 5.4 is the space photo, the image after image segmentation, and the image after the two were superimposed in the software Photoshop.

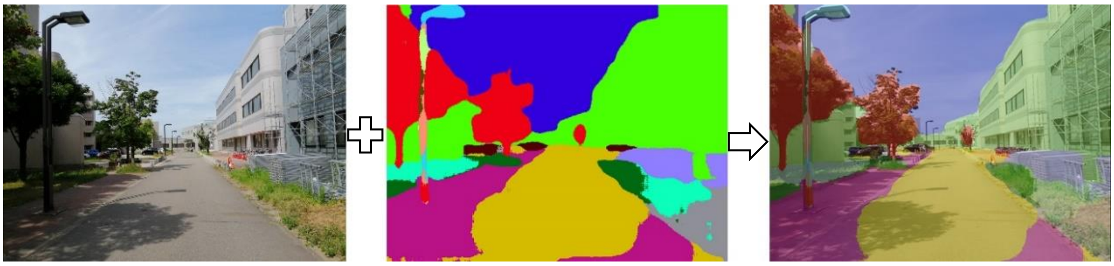


Figure 5.4 Overlay of the original image and the image after semantic segmentation.

Figure 5.4 shows a good coincidence between the segmented image object (color blocks) and the original image objects, which shows that the data of object area proportion obtained by the software is effective. Then, we did semantic segmentation for each photo and got the data set of 13 objects of all 433 photos (because two adjacent photos have similar image content, one photo was retained for every two photos. After image reduction, 586 photos remain. Then, the photos with unclear images are deleted, and 433 photos remain).

5.2.2 Two-step clustering of spatial features

Data preprocessing

First, we combine physical features and image features. The photos were taken every ten steps, and the physical features are the average value of each section, so we use the image features as the basis and merge the physical features according to the section where the image is located. We got a data set of images and physical features of 433 samples.

Second, 23 features are reduced by principal component analysis to remove the features with low weight. The principal component analysis is the most widely used method, and it is necessary to use the categorical principal components analysis (CATPCA) when there are both continuous and categorical variables. Categorical principal components analysis (CATPCA) is appropriate for data reduction when variables are categorical (e.g., ordinal) and the researcher is concerned with identifying the underlying components of a set of variables (or items) while maximizing the amount of variance accounted for in those items (by the principal components). The main advantage of using CATPCA instead of traditional PCA is to model the nonlinear relationship between variables.

Since the purpose of this principal component analysis is not to reduce dimensionality but to calculate weights, the higher the interpretation of the cumulative variance, the better, so we try to set the number of principal components to 10 and perform dimensionality reduction-optimal scaling analysis (Optimal Scaling).

However, after performing the optimal scale analysis and using the entropy weight method (EWM) to calculate the weights, it is found that some weight values are negative. Two reasons cause this. One is that the default data standardization of SPSS is unsuitable for the data set, and the EWM needs to be used. Data standardization is performed, and variables are divided into positive and negative variables. After normalization, all the values that are 0 in the data set are changed to 0,01 to meet the requirements of CATPCA. Finally, we set the number of primary components to ten, perform optimal scaling analysis (dimension reduction-optimal scaling analysis) and calculate the weights of variables (Table 5.1).

Six features with small weights were removed through calculation, and 17 main features remain (Table 5.2).

Table 5.1 Weights of 23 variables (physical and image features)

Features	weights
Id_5tree_10grass_18plant	6.47%
W2H2	6.11%
Id_1wall-2building_26house	6.08%
Id_14earth	5.90%
Width between boundary trees (W2) (m)	5.64%
Width of soft boundary	5.36%
Number of material types	5.15%
Id_7road_12sidewalk_53path_54stairs	5.05%
Id_21car	4.81%
W1H1	4.54%
Height of building boundary (H1) (m)	4.48%
Number of elements	4.46%
Height of boundary tree (H2) (m)	4.31%
Id_88streetlight	4.30%
Number of space types	3.78%
Id_22water_61river	3.73%
Id_3sky	3.64%
Percentage of main space length	3.47%
Number of boundary layers	3.43%
Id_35rock	3.26%
Width between pavement boundary (W3) (m)	3.06%
Width between buildings (W1) (m)	1.87%
Id_33fence	1.11%

Third, there are continuous variables and ordinal variables in the data set. If the ordinal variables are used directly, the importance of ordinal variables in clustering may be emphasized too much. So, we convert categorical variables in features into dummy variables. Four categorical variables were among the 17 feature variables and were transformed into 24 dummy variables according to the number of categorical variables.

Two-step clustering

Cluster analysis is generally divided into three categories, namely two-step clustering, K-means clustering, and hierarchical clustering. Since the K-means clustering algorithm cannot analyze categorical variables, systematic clustering cannot judge the

Table 5.2 Seventeen main features after dimension reduction.

Number	Spatial physical features （13）	Proportion features of image objects （4）
1	Width between buildings (W_1) (m)	Id_1wall2building_26house
2	Width between boundary trees (W_2) (m)	Id_3sky
3	Width between pavement's boundaries (W_3) (m)	Id_5tree_10grass_18plant
4	Height of building's boundary(H1) (m)	Id_7road_12sidewalk_53path_54stairs
5	Height of boundary tree(H_2) (m)	
6	W_1/H_1	
7	W_2/H_2	
8	Width of soft boundary	
9	Percentage of main space length	
10	Number of boundary layers	
11	Number of space types	
12	Number of material types	
13	Number of elements	

quality of clustering. So, we chose the two-step clustering algorithm as a clustering algorithm used in SPSS Modeler, an improved version of the BIRCH hierarchical clustering algorithm. It can be applied to clustering mixed-attribute datasets and added a mechanism to automatically determine the optimal number of clusters, making the method more practical. The two-step clustering algorithm is divided into two stages: 1) Pre-clustering stage. The idea of Cluster Feature Tree（CF tree）growth in the BIRCH algorithm is adopted, and the data points in the data set are read individually. While generating the CF tree, the data points in the dense area are pre-clustered to form many small sub-clusters. 2) Clustering stage. The result subclusters in the pre-clustering stage are taken as the object. The agglomerative hierarchical clustering method is adopted to merge the subclusters one by one until reaching the required number of clusters.

The two-step clustering method of unsupervised learning was used for spatial classification. The value range of the features of high emotional stimulation quality space and low emotional stimulation quality space was determined according to the spatial value.

First, determine the proportion of outliers in the data set. To determine the reasonable proportion of outliers, in SPSS software, we assume that the outliers are 0%, 5%, 15%, and 25% and perform a two-step clustering analysis on the data sets of image and physical features. The results show that the classification is most reasonable when the upper outlier is 5%, and the separation results are consistent with the actual space situation. According to the final output result, the system automatically removes 8.3%, that is, 36 outliers and the number of samples that can be used for clustering is 397.

Second, two-step clustering is performed to obtain the final number of clusters.

Third, according to the clustering results, draw a box diagram of feature distribution, and calculate the quartile value (maximum, upper quartile (Q3), lower quartile (Q1), and minimum and median (Q2) of each feature in each cluster). The median (Q2) can reflect the average level of a group of data, and Q3-Q1 can reflect the dispersion of data (Appendix G explains the quartile values of the box diagram).

Fourth, to compare the valency value of each section, find the sections with valency values > 1 and < 0; that is, to find the space with high emotional stimulation quality and the space with low emotional stimulation quality. Then, we compared the median and Q3-Q1 values of the sections corresponding to these two types of spaces, drew a broken line diagram, and then found the features with apparent differences in high and low-quality spaces through visual words. Hence, we extracted the value range or quantitative description of the features that affect the emotion-eliciting quality of the space.

Fifth, to draw ten spatial distribution maps in the one-dimensional coordinate system with emotional value as the axis and evaluate the spatial quality and features according to the spatial function and the difference between China and Japan.

5.3 Results

5.3.1 Association between spatial features and emotional-eliciting quality

1) This study extracted ten image features and 13 physical space features from 10 public spaces. Among the image features, the proportion of 6 is tiny, so the four main

image features were retained. Principal component analysis was used to reduce the feature dimension and calculate each feature's weight (Table 5.3).

Among the 23 features, the weight of the highest feature is 6.47%, and the lowest is 1,11%. Among them are 8 with a weight greater than 5%, 14 with a weight greater than 4%, and 21 with a weight greater than 3%. These features all have a specific impact on spatial emotions. The proportion of plants, the width and height ratio of space formed by trees, and the proportion of buildings impact spatial emotions more than other features.

Table 5.3 Weight of seventeen main features.

Features	weights
Id_5tree_10grass_18plant	6.47%
W2/H2	6.11%
Id_1wall-2building_26house	6.08%
Width between boundary trees (W2) (m)	5.64%
Width of soft boundary (m)	5.36%
Number of material types	5.15%
Id_7road_12sidewalk_53path_54stairs	5.05%
W1/H1	4.54%
Height of building boundary (H1) (m)	4.48%
Number of elements	4.46%
Height of boundary tree (H2) (m)	4.31%
Number of space types	3.78%
Id_3sky	3.64%
Percentage of main space length	3.47%
Number of boundary layers	3.43%
Width between pavement boundaries (W3) (m)	3.06%
Width between buildings (W1) (m)	1.87%

By comparing the clustering results with different proportions of outliers, it was found that when 5% of the outliers were removed, the clustering results could distinguish different sections in the space, which was in line with the actual spatial features. In the two-step clustering analysis output (Tables 5.4, 5.5), the recommended optimal clustering results were five categories, and the Silhouette coefficient was 0.3132 within a fair range (Figure 5.5), so we accepted the results and got the clustering result of each sample and the quartile values of 17 features in the five clusters. (Tables 5.6).

Table 5.4 Output of the two-step clustering: description of ten kinds of clustering.

Auto-Clustering

Number of Clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change[a]	Ratio of BIC Changes[b]	Ratio of Distance Measures[c]
1	-15229.989			
2	-16737.290	-1507.301	1.000	1.347
3	-17759.491	-1022.201	.678	1.158
4	-18590.779	-831.287	.552	1.498
5	-19020.319	-429.540	.285	1.482
6	-19187.820	-167.501	.111	1.000
7	-19355.230	-167.411	.111	1.180
8	-19439.687	-84.456	.056	1.072
9	-19493.304	-53.617	.036	1.160
10	-19487.660	5.644	-.004	1.014

Table 5.5 Output of the two-step clustering: distribution of recommended clusters divided into five categories.

Cluster Distribution

		N	% of Combined	% of Total
Cluster	1	95	21.9%	21.9%
	2	51	11.8%	11.8%
	3	74	17.1%	17.1%
	4	100	23.1%	23.1%
	5	77	17.8%	17.8%
	Outlier (-1)	36	8.3%	8.3%
	Combined	433	100.0%	100.0%
Total		433		100.0%

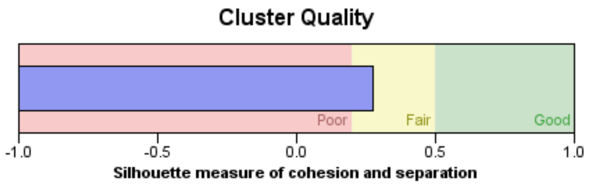


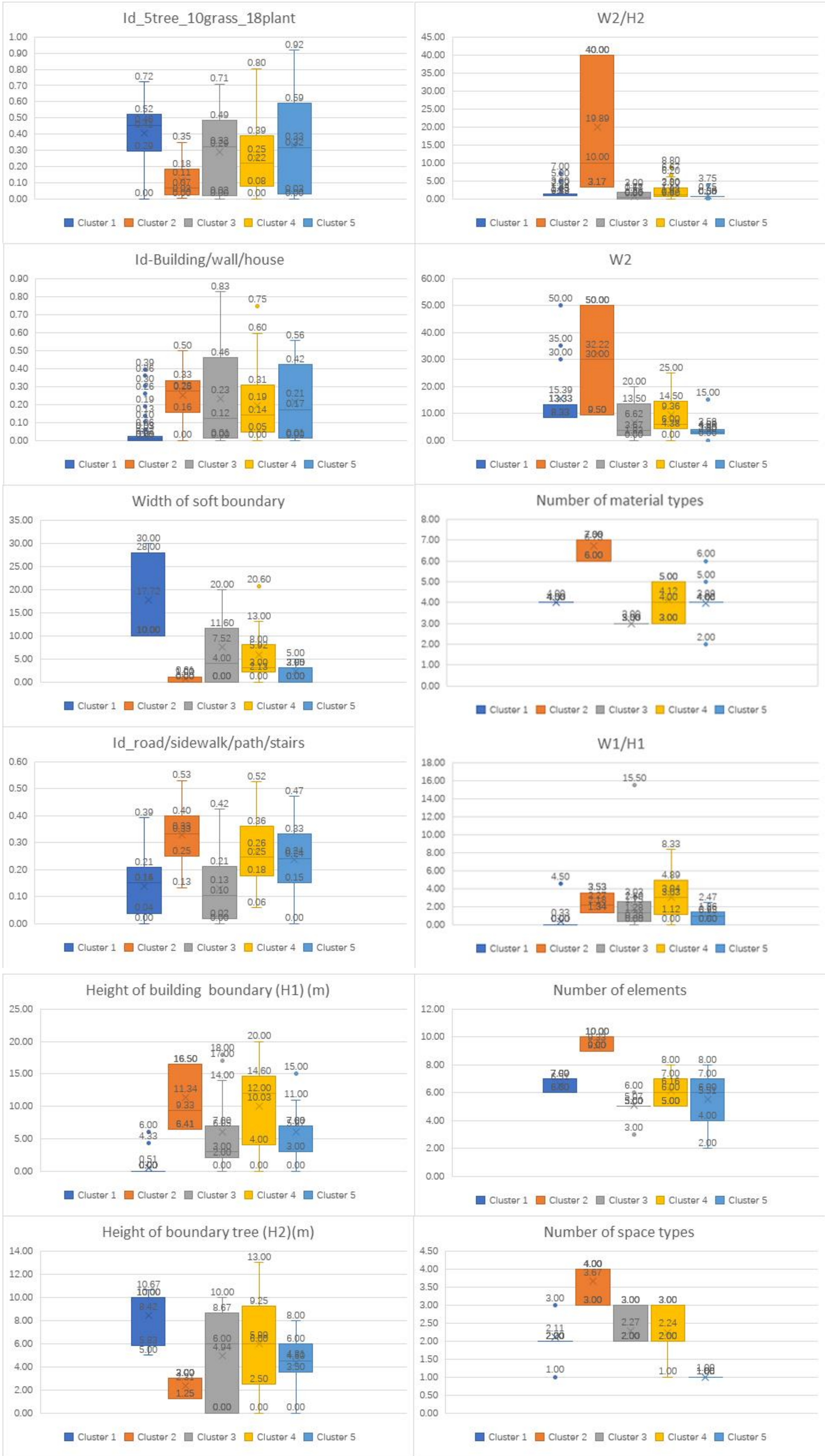
Figure 5.5 Output of the two-step clustering: model summary and cluster quality (the Silhouette coefficient is 0.3132 within a fair range).

Table 5.6 The quartile values of 17 features in the five clusters.

Features	Quartile of cluster 1			Quartile of cluster 2			Quartile of cluste 3			Quartile of cluste 4			Quartile of cluste 5			Quartile of Overall		
	25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%
Id_5tree_10grass_18plant	0.30	0.46	0.52	0.03	0.07	0.18	0.02	0.32	0.49	0.08	0.22	0.39	0.03	0.32	0.59	0.07	0.27	0.49
W2/H2	0.89	1.30	1.49	3.20	10.07	39.93	0.08	0.40	1,71	0.62	1.02	3.04	N/A	0.55	N/A	0.51	0.90	3.06
Id_1wall-2building_26house	0.00	0.01	0.02	0.16	0.28	0.33	0.01	0.12	0.45	0.05	0.14	0,31	0.01	0.17	0.42	0.01	0.09	0.30
Width between boundary trees W2 (m)	8.31	8.36	13.35	9.62	30.09	49.92	1.77	3.68	13.53	5.17	6.10	14.53	2.52	3.05	4.02	3.66	8.28	14.57
Width of soft boundary (m)	10.01	10.04	28.02	0.05	0.99	1.02	0.05	4.04	11,61	2.37	3.07	7.95	0.07	3.01	3.06	1.01	3.08	10.04
Number of material types	N/A	4	N/A	N/A	7	N/A	N/A	3	N/A	N/A	N/A	N/A	N/A	4	N/A	N/A	N/A	N/A
Id_7road_12sidewalk_53path_54stairs	0.04	0.15	0.21	0.25	0.33	0.40	0.02	0.10	0.21	0.18	0.25	0.36	0.16	0.24	0.33	0.11	0.21	0.31
W1/H1	N/A	0.02	N/A	1.35	2.19	3.53	0.36	1.26	2.49	1.10	3.05	4.89	0.02	0.94	1.36	0.02	1.09	2.48
Height of buildings boundary H1 (m)	N/A	0.03	N/A	6.45	9.34	16.52	2.01	3.03	7.04	6.01	12.03	14.60	3.02	7.01	7.05	0.04	6.41	12.01
Number of elements	N/A	7	N/A	N/A	9	N/A	N/A	5	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Height of boundary tree H2 (m)	5.83	9.99	10.01	1.26	3.00	3.01	0.03	6.00	8.65	2.49	5.99	8.42	3.48	4.50	6.00	3.00	5.84	8.65
Number of space types	N/A	2	N/A	N/A	4	N/A	N/A	2	N/A	N/A	2	N/A	N/A	1	N/A	N/A	N/A	N/A
Id_3sky	0.04	0.11	0.23	0.13	0.18	0.26	0.00	0.03	0.06	0.06	0.14	0.21	0.04	0.08	0.13	0.04	0.10	0.18
Percentage of main space length	0.55	0.76	0.78	0.55	0.56	0.57	0.55	0.56	0.63	0.69	0.77	0.86	0.95	1.00	1.00	0.56	0.76	0.87
Number of boundary layers	N/A	2	N/A	N/A	1.5	N/A	N/A	1	N/A	N/A	2	N/A	N/A	3	N/A	N/A	N/A	N/A
Width between pavement boundaries W3 (m)	4.99	4.99	5.03	4.03	4,45	6.75	2.04	2.70	5.01	3.99	4,99	6.53	1.97	3.28	7.98	2.70	4.48	5.05
Width between buildings W1 (m)	N/A	0.21	N/A	12.67	22.74	36.12	2.71	4.11	5.19	7.63	28.61	53.70	0.21	6.45	12.16	0.24	5.16	23.04

Next, according to the output of the two-step clustering (Appendix F), we drew a box diagram of each variable to obtain the feature variables' median value and value range (Figure 5.7). The results show that almost all spatial features have overlapping parts between different clusters, which means that the boundaries between these features are unclear. It is necessary to use multiple features to work together to determine that the sample belongs to which cluster. Since it is difficult to find the difference between clusters from each feature, we choose the space corresponding to valence>1 as the space of high valence, the space corresponding to valence<0 as the space of low valence, and then look for the emotional valence and the corresponding space feature relationship.

Then, according to the output of the two-step clustering analysis, each sample was classified into 5 clusters, and we got clustering results for each spatial section. The table shows the classification results of each section after a two-step clustering analysis (Table 5.7)


Id_5tree_10grass_18plant
W2/H2
Id-Building/wall/house
W2
Width of soft boundary
Number of material types
Id_road/sidewalk/path/stairs
W1/H1
Height of building boundary (H1) (m)
Number of elements
Height of boundary tree (H2)(m)
Number of space types
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5

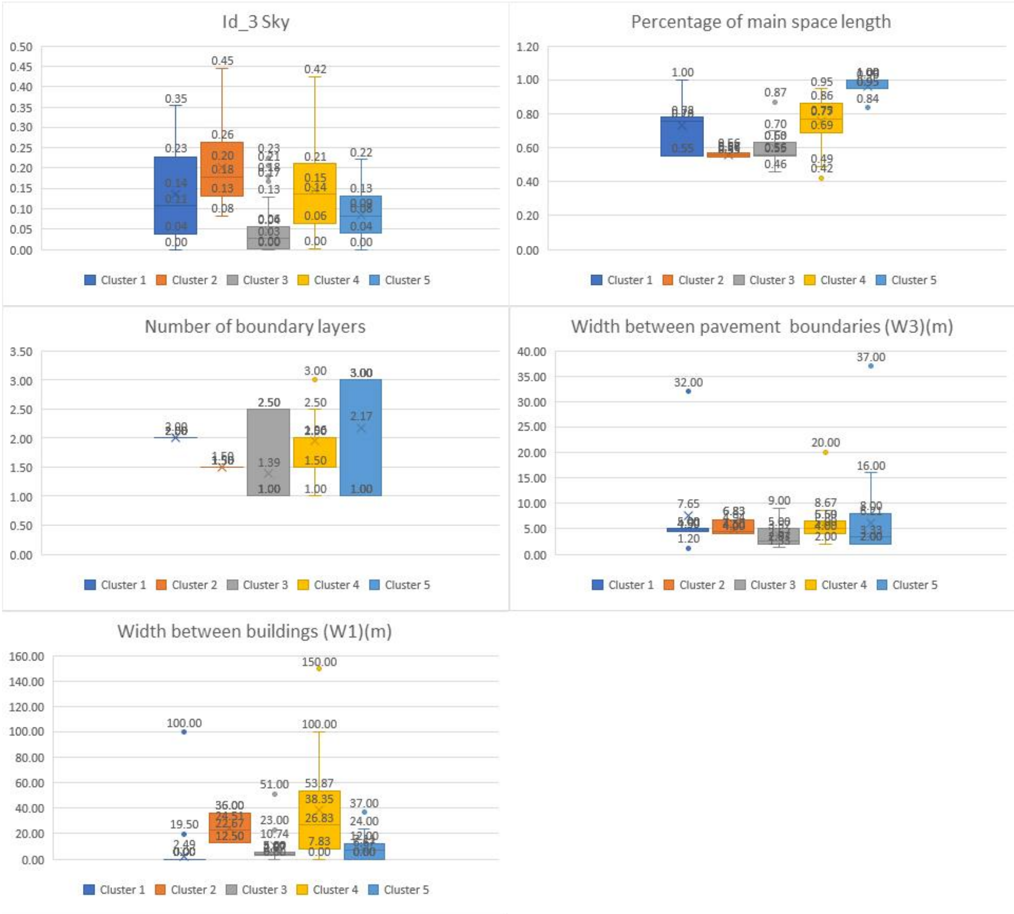


Figure 5.6 Box diagram of spatial feature data distribution.

Since the dimensions of the 17 features are inconsistent, we divided the features into two categories: features with values between zero and one and features greater than one, and drew the feature box graph with low and high valence, respectively (Figures 5.7, 5.8, 5.9, and 5.10).

From the boxplot for each feature, we extracted the upper, upper quartile, median, mean, lower quartile, and lower values of the feature's boxplot.

From the boxplot of each feature, we extracted the upper, upper quartile, median, mean, lower quartile, and lower values of the feature's boxplot and calculated the value of Q3-Q1. The median can reflect the average level of a set of data, and Q3-Q1 can reflect the degree of dispersion of the data.

Then we compare the median and Q3-Q1 of the spatial eigenvalues with high and low valence using line charts (Figure 5.11, 5.12, 5.13, and 5.14).

Table 5.7 Classification results of each section after a two-step clustering analysis.

Sites	Valence	Five Clusters（5% Outlier Removed）
Huarun twenty-four city park-s3	1.64	1
Residential area in Yokaichi-s3	1.11	5
D. T. Suzuki Museum-s3	1.11	3
DLPU campus-s3	1.09	3
Dalian ganjingli· dongshi District-s3	1.09	4
Huarun twenty-four city park-s2	0.91	4
JAIST campus-s2	0.89	4
Higashi Chaya District-s3	0.89	3
JAIST campus-s1	0.78	4
Kenroku-en-s3	0.78	1
D. T. Suzuki Museum-s2	0.78	1
Dalian heroes memorial park-s4	0.73	5
Higashi Chaya District-s4	0.67	4
Kenroku-en-s1	0.67	3
Kenroku-en-s2	0.67	1
JAIST campus-s4	0.56	4
Kenroku-en-s4	0.56	1
D. T. Suzuki Museum-s4	0.56	1
DLPU campus-s1	0.55	3
Huarun twenty-four city park-s4	0.45	4
Residential area in Yokaichi-s1	0.44	4
Higashi Chaya District-s1	0.44	4
Higashi Chaya District-s2	0.44	3
Dalian heroes memorial park-s2	0.36	5
Residential area in Yokaichi-s2	0.33	2
Residential area in Yokaichi-s4	0.33	2
Meilin park residential area-s2	0.30	4
Dalian heroes memorial park-s3	0.27	5
Huarun twenty-four city park-s1	0.18	N/A
Dalian ganjingli· dongshi District-s2	0.18	5
JAIST campus-s3	0.11	2
D. T. Suzuki Museum-s1	0.11	3
Dalian ganjingli· dongshi District-s1	0.09	5
DLPU campus-s2	-0.09	3
DLPU campus-s4	-0.18	N/A
Dalian ganjingli· dongshi District-s4	-0.18	5
Dalian heroes memorial park-s1	-0.27	4
Meilin park residential area-s3	-0.52	5
Meilin park residential area-s4	-0.53	5
Meilin park residential area-s1	-0.70	4

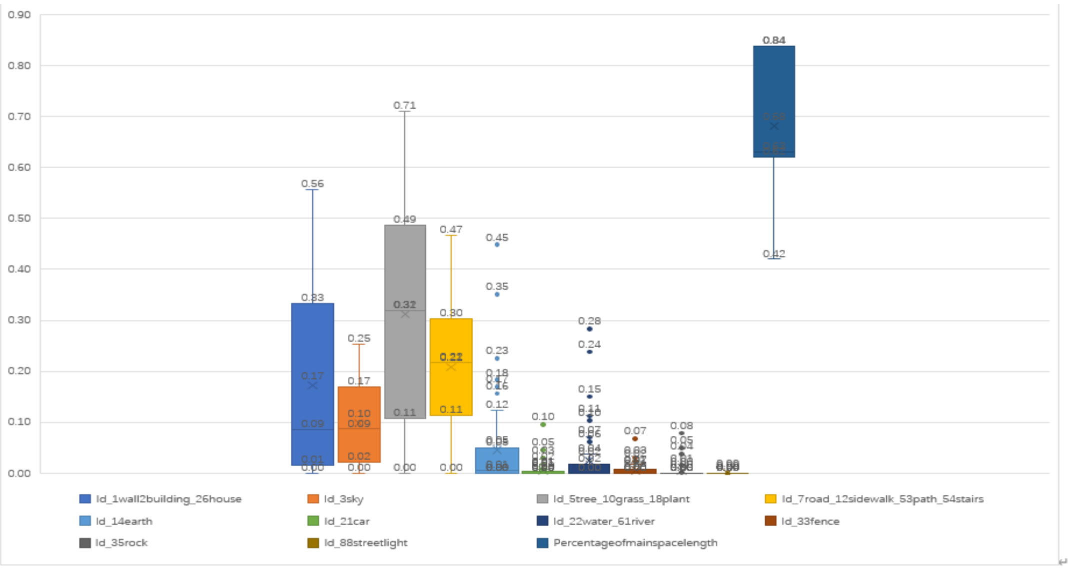


Figure 5.7 Spatial features of high emotional valence (features with valence between zero and one).

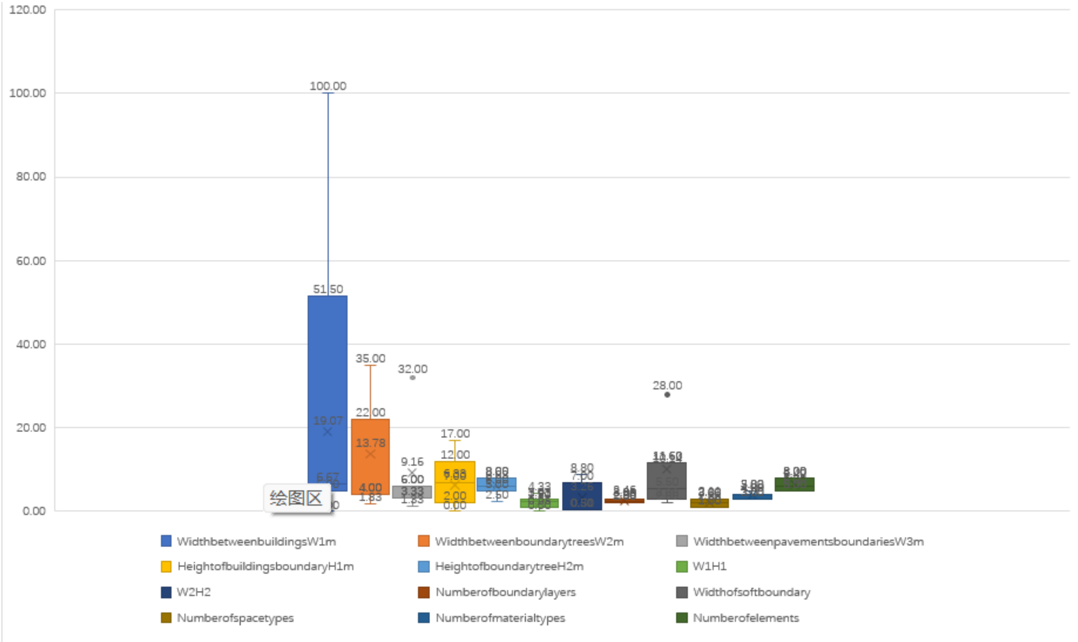


Figure 5.8 Spatial features of high emotional valence (features with a valence greater than one).

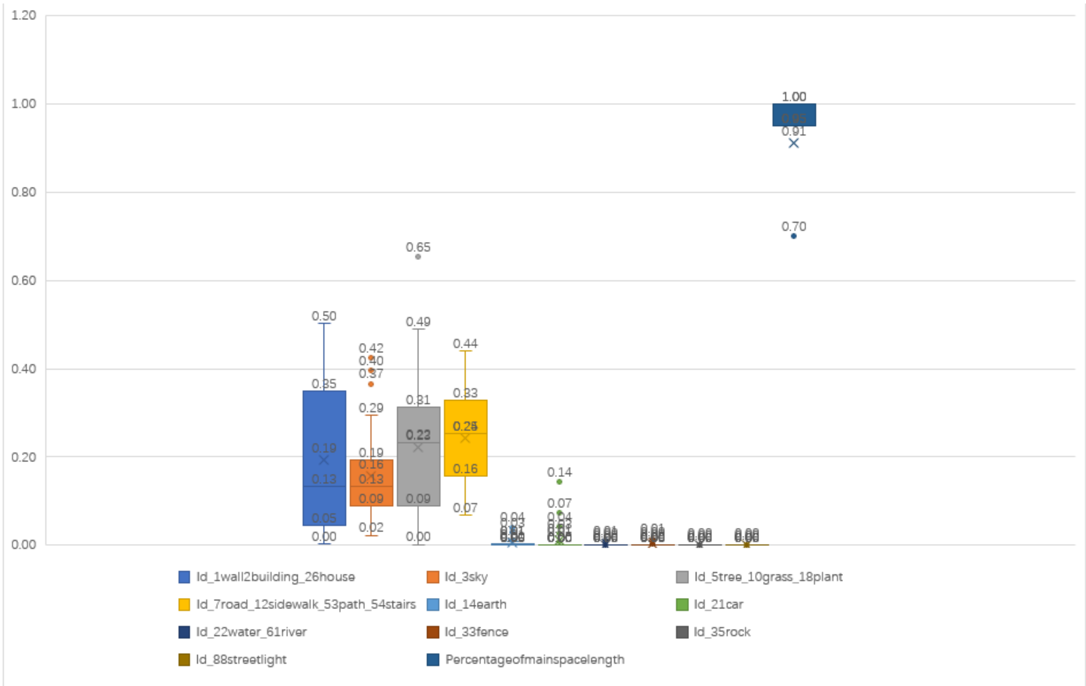


Figure 5.9 Spatial features of low emotional valence (with a valence between zero and one).

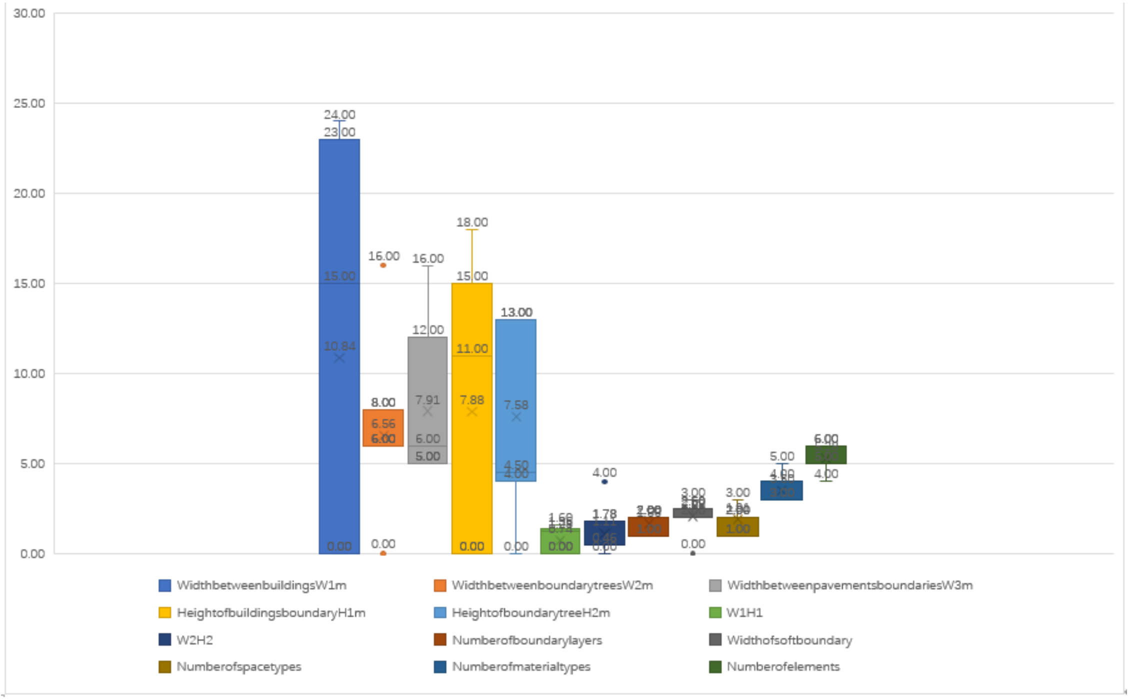


Figure 5.10 Spatial features of low emotional valence (with a valence greater than one).

We divide the feature data into the continuous and the categorical variable data set and conduct the Independent-samples t-test for the continuous variable data set and the Chi-square test for the categorical variable data set. Because the continuous variable data did not meet the normal distribution, we used the Mann-Whitney U test. Then we got the Asymptotic Significance (2-sided) from the output of the tests; when the significance is less than 0.05 ($P<0.05$), the difference between the high and low valence is significant (Table 5.8).

The results show that the difference of the nine features between the high and low valence is significant: Id_5tree_10grass_18plant, W2/H2, Width between boundary trees (W2) (m), Width of soft boundary (m), W1/H1, Percentage of main space length, Number of boundary layers, Width between pavement boundaries (W3) (m), Width between buildings (W1) (m). So these nine features might distinguish the high or low emotional valence. Furthermore, these nine main features work together to distinguish whether the space is positive or negative, and the remaining features play an auxiliary role (Table 5.9).

Table 5.8 Comparison of the median, mean, and Q3-Q1 with high and low valence.

Features	High emotional valence			Low emotional valence			Asymptotic Significance (2-sided) of median difference
	Median (H)	Mean (H)	Q3-Q1 (H)	Median (L)	Mean (L)	Q3-Q1 (L)	
Id_5tree_10grass_18plant	32.00%	31.00%	38.00%	21.00%	22.00%	22.00%	0.048*
W2/H2	3.26	3.26	6.50	1.70	1.11	1.32	0.002**
Id_1wall-2building_26house	9.00%	17.00%	32.00%	13.00%	19.00%	28.00%	0.256
Width between boundary trees (W2) (m)	22.00	13.78	18.00	6.00	6.56	2.00	0.036*
Width of soft boundary (m)	5.50	10.14	8.60	2.50	2.08	0.50	0.000***
Number of material types	3.00	3.82	1.00	N/A	3.50	1.00	0.056
Id_7road_12sidewalk_53 path_54stairs	22.00%	21.00%	19.00%	25.00%	24.00%	18.00%	0.235
W1/H1	1.91	1.91	2.08	1.28	0.74	1.36	0.002**
Height of buildings boundary (H1) (m)	7.00	6.33	10.00	11.00	7.88	15.00	0.833
Number of elements	6.00	6.40	2.00	N/A	5.28	1.00	0.062
Height of boundary tree (H2) (m)	6.00	6.03	3.00	4.50	7.58	9.00	0.694
Number of space types	2.00	2.11	2.00	N/A	1.91	1.00	0.253
Id_3sky	9.00%	10.00%	15.00%	13.00%	16.00%	10.00%	0.082
Percentage of main space length	63.00%	68.00%	22.00%	95.00%	91.00%	14.00%	0.000***
Number of boundary layers	2.50	2.46	1.00	N/A	1.66	1.00	0.000***
Width between pavement boundaries (W3) (m)	3.33	3.45	2.67	6.00	7.91	7.00	0.002**
Width between buildings (W1) (m)	6.67	19.07	46.50	15.00	10.84	23.00	0.033*

*Note: Q3 is the upper quartile, and Q1 is the lower quartile. * P<0.5; ** P<0.01; *** P<0.001.*

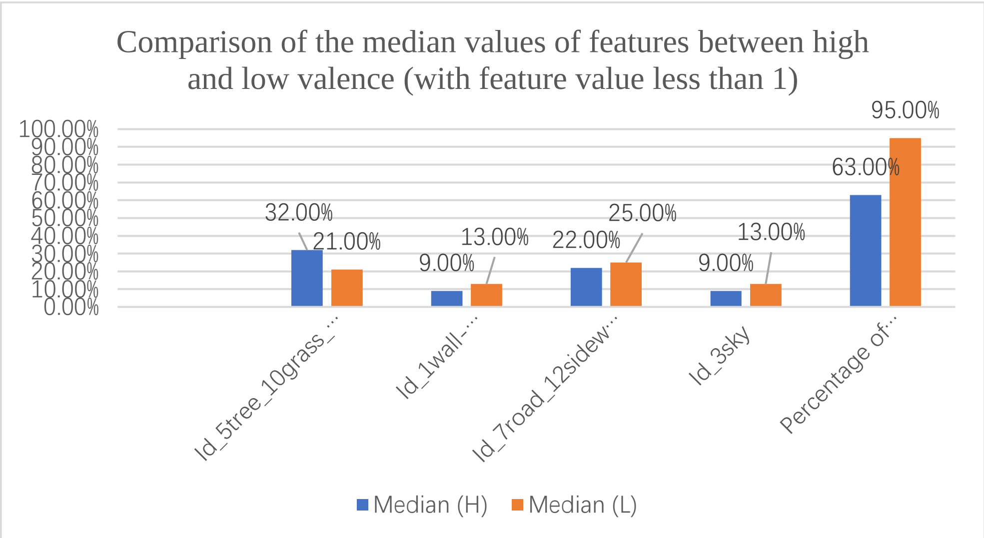


Figure 5.11 Comparison of the median values of features between high (H) and low (L) valences (feature value is less than one).

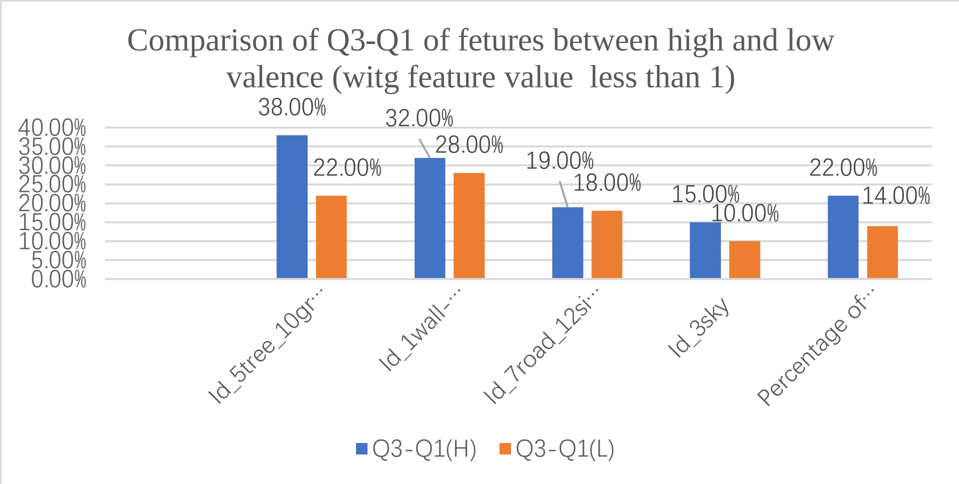


Figure 5.12 Comparison of Q3-Q1 of features between high (H) and low (L) valences (feature value is less than one).

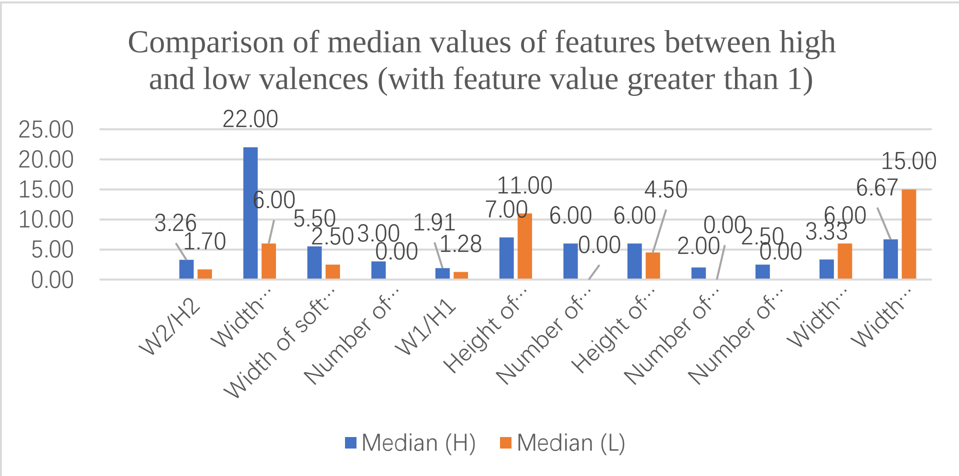


Figure 5.13 Comparison of the median of features between high (H) and low (L) valences (feature value is greater than one).

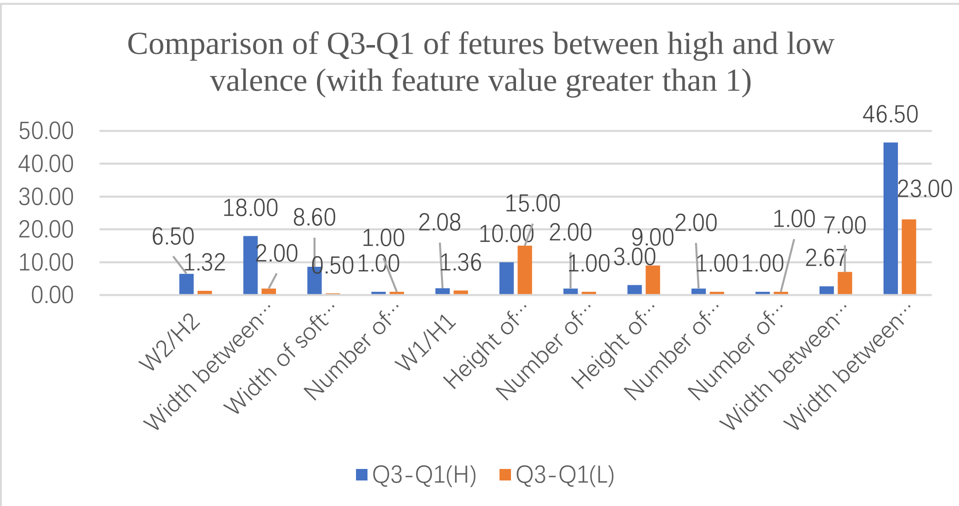


Figure 5.14 Comparison of Q3-Q1 of features between high (H) and low (L) valences (feature value is greater than one).

Table 5.9 Comparison of spatial feature value ranges of high and low emotion-eliciting quality.

Features	High emotional quality			Low emotional quality		
	Upper-quartile (Q3)	Median (H)	Lower-quartile (Q1)	Upper-quartile (Q3)	Median (L)	Lower-quartile (Q1)
Id_5tree_10grass_18plant	49.00%	32.00%	11.00%	31.00%	21.00%	9.00%
W2/H2	7.00	3.26	0.50	1.78	1.70	0.46
Width between boundary trees (W2) (m)	22.00	17.50	4.00	8.00	6.00	6.00
Width of soft boundary (m)	11.60	5.50	3.00	2.50	2.50	2.00
W1/H1	3.03	1.91	0.95	1.36	1.28	0.00
Percentage of main space length	84.00%	63.00%	62.00%	100.00%	95.00%	86.00%
Number of boundary layers	3.00	2.50	2.00	2.00	1.32	1.00
Width between pavement boundaries (W3) (m)	6.00	3.33	3.33	12.00	6.00	5.00
Width between buildings (W1) (m)	51.50	6.67	5.00	23.00	15.00	0.00

5.3.2 Basic features of positive and negative spaces

The following spatial indicators might be used to evaluate the spatial emotion-eliciting quality (need to be verified in practice).

Features of public space with high emotion-eliciting quality are as follows:

a. The proportion of visual images of trees is between 11% and 49%, with a median of 32%;

b. The W2/H2 is between 0.5 and 7.00, with a median value of 3.26;

c. The width between boundary trees (W2) (m) is between 4m and 22m, and the median value is 17.5m.

d. The width of the soft boundary (m) is between 3M and 11.6m, and the median value is 5.5m.

e. W1/H1 is between 0.95 and 3.03, with a median of 1.91.

f. The Percentage of main space length is between 62% and 84%, with a median of 63%;

g. The number of boundary layers is between 2 and 3, and the median value is 2.5;

h. The width between payment boundaries (W3) (m) is between 3.33M and 6.0m, and the median value is 3.5m;

i. The width between buildings (W1) (m) is between 5m and 51.5m, and the median value is 6.67m.

Features of public space with low emotion-eliciting quality are as follows:

a. The proportion of visual images of trees is between 9% and 31%, with a median of 21%;

b. The W2/H2 is between 0.46 and 1.78, with a median value of 1.7;

c. The width between boundary trees (W2) (m) is between 6m and 8m, and the median value is 6.5m.

d. The width of the soft boundary (m) is between 2M and 2.5m, and the median value is 2.6m.

e. The W1/H1 is between 0 and 1.36, with a median of 1.28

f. The percentage of main space length is between 86% and 100%, with a median of 95%;

g. The number of boundary layers is between 1 and 2, and the median value is 1.32;

h. The width between payment boundaries (W3) (m) is between 5M and 12M, and the median value is 6m;

i. The width between buildings (W1) (m) is between 0 and 23m, and the median value is 15m.

In addition, we made the following Figure 5.15 according to the average value of the emotional valence of other spaces. The Y-axis of the figure is emotional valence, and the X-axis is the color block graph for image semantic segmentation of five types of public spaces. This figure shows the difference and distribution of the features of each type of space based on valence.

5.4 Discussion

1) The nine main features jointly affect the emotional-eliciting quality of space.

The distribution of weights of features is relatively uniform. Among the 23 features, the weight of the highest feature is 6.47%, and the lowest is 1,11%. Among them are 8 with a weight greater than 5%, 14 with a weight greater than 4%, and 21 with a weight greater than 3%. These features all have a particular impact on spatial emotion. Among them, the proportion of plants, the aspect ratio of the space formed by plants, and the proportion of buildings impact the spatial mood more than other features.

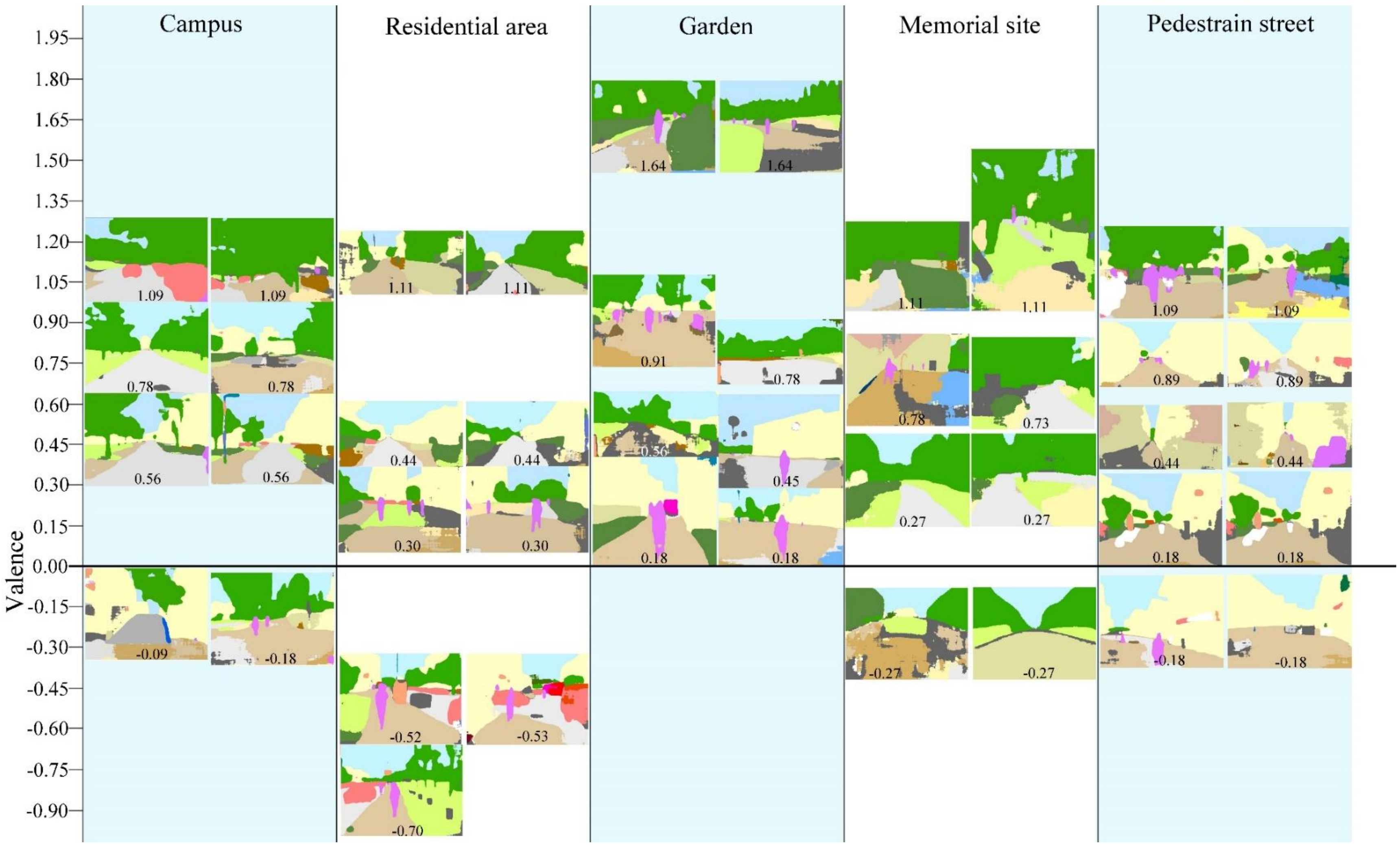


Figure 5.15 Distribution diagram of all ten public spaces with the axis of valence

The proportion of the space and the distance between the buildings on both sides have little influence on user emotions. This result shows that the nine main features jointly affect the emotional-eliciting quality of space, and it is necessary to combine the range and median value of these features to judge the emotional-eliciting quality of space comprehensively.

2) The change range of emotional valence of users in space is related to the spatial attribute.

The color patch image semantic segmented was obtained using the convolutional neural network technology to extract objects in the spatial image. A schematic diagram of the distribution of all ten images was established with valence as the axis: From the above figure, it can be found:

a. Among all the pictures, the user has the highest emotional valence in the park space, and there is no space with a valence<0. The remaining four types of spaces have space segments with valence < 0; this shows that users have the best emotional response in the park space.

b. From high valence to low valence, the proportion of trees in the image tends to decrease, and the ratio of buildings increases.

c. Among the five types of spaces, the emotional effect value of the residential area has the most considerable change range, which is 1.11-(-0.70)=1.81, which indicates that the spatial emotion-eliciting quality of the residential area has the most noticeable change, that is, the consistency is poor. In contrast, the campus and pedestrian streets have a minor change range. The range of emotional effect value of space is 1.09-(-0.09)=1.18 and 1,09-(-0.18)=1.27, respectively, which indicates that the change of spatial emotion-eliciting quality is small; that is, the spatial consistency is good.

3) Compared with previous studies, this study proposed the spatial features that affected users' emotions and put forward the quantitative value range of positive and negative space features.

Some results of this research support previous research, while some results do not. Among the ‘scale’ indexes, the results regarding the two indexes related to user emotions, 'width between pavement boundaries (W_3)' and 'height of boundary trees (H_2)', support that of Lee et al. (2009) and Bivina et al. (2018). The results regarding the two indexes of 'spatial continuity' and 'continuity of spatial boundary' are similar to the indexes of 'continuity' and 'continuity of wall street' proposed by Bivina et al. (2018) and Tang et al. (2019). However, for the indicators 'width between buildings (W_1)', 'height of building on both sides (H_1)', 'W_1/H_1', 'sky ratio', 'space visual entropy', and 'enclosure degree', this study's results contradict those of some previous studies (Schneider et al. 2014; Jahanmohan, T. 2016; Li et al. 2016; Tang et al. 2019). Different data analysis methods may cause these differences.

In addition to the spatial features that affect users' emotions, this study further proposes the quantitative value range of the features of positive and negative space, which has not been found in previous studies.

5.5 Summary

In this chapter, FCN algorithm software is used to perform semantic segmentation on the spatial image, extract the image features of the proportion of spatial elements, and form the spatial feature dataset with the physical features extracted in the actual space. Then, five clusters are extracted using the two-step clustering method (unsupervised learning). The data box chart analysis reveals the differences between various types. Finally, combined with the average emotional valence of space, we obtained the main features corresponding to high and low valence. The extensive use of these quantitative features with a range of values might meet the practice's needs and support the design of urban public space, the renewal of old space, and the managemental decision of urban space.

Chapter 6

Comparison of the features of public spaces between Japan and China

6.1 Introduction

This chapter aims at sub-goal 3: the similarity and differences in the features of Japanese and Chinese public spaces.

Comparing public space between the two countries mainly includes comparing environmental features and cultural backgrounds. This chapter conducted a comprehensive comparative analysis of Japanese and Chinese public space based on physical, image, and user perception features. The results of this chapter will help us understand the similarities and differences between Japanese and Chinese public spaces, as well as that among each type of space.

6.2 Methods

6.2.1 Weights of perception features

Applying the principal component analysis (PCA) and the entropy weight method (EWM) can calculate each factor's weight and reduce the dimensionality of the data. The steps of factor analysis are as follows:

1) Data cleaning to remove missing items and samples with too large deviations in the data;

2) Kaiser-Meyer-Olkin (KMO) test and Bartlett's Test

First, the Statistical Product and Service Solutions (SPSS) will be used to test the validity of the data, that is, the Kaiser-Meyer-Olkin (KMO) and Bartlett's test of Sphericity. The KMO values closer to 1.0 are considered ideal, while values less than 0.5 are unacceptable, and the result of Bartlett's test of Sphericity shows that Sig.<0.05

means that the data meets the standard and each variable is independent. So, if the validity test meets the requirements, factor analysis can be conducted.

3) Principal component analysis (PCA)

Then, we can conduct a PCA on the 12 features (data from the questionnaire in Table 3.3) using SPSS and get two tables from the output. One is the table of Total variance explained; the other is the table of the Rotated component matrix. If the cumulative contribution of common factors with an initial eigenvalue greater than 1 is more outstanding than 80%, the common factors have a reasonable interpretation of all factors.

4) Calculating the weights of all features

We will use the entropy weight method to calculate the weight of each feature. The method is the same as the method introduced in chapter 4 (P58). Once obtaining the weights of all features, we decrease the number of features according to it.

6.2.2 Comparative analysis of spatial features

In Chapter 5, we extracted 17 main physical and image features of public spaces that affect users' emotions (Table 5.3). After merging the physical, image, and perception features, we will compare the features of public spaces between Japan and China.

First, we use SPSS to calculate the average value to fill in the missing values of features in the dataset.

Second, we analyze the significance of the difference in all the features between Japan and China. We divide the feature data into the continuous variable data set, and the categorical variable data set and conduct the Independent-samples t-test for the continuous variable data set and the Chi-square test for the categorical variable data set. For the Independent-samples t-test, the data set must meet normally distributed. So, we use descriptive statistics in SPSS to test whether the continuous variable data meet the normal distribution. If not, we will use the Mann-Whitney U test. Then we will get the Asymptotic Significance (2-sided) from the output of the test; when the significance is less than 0.05 ($P<0.05$), the difference between Japan and China is noticeable; on the contrary, they are similar. According to the same method, we can find the difference and similarities between each type of public space between Japan and China. In addition to the different significance analyses, we also choose the features' median to analyze the difference's specific value.

Third, drawing bar charts to visualize the similarities and differences between the two countries and analyzing the results.

6.3 Results

6.3.1 Main spatial features

The results of the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of Sphericity show that the KMO value is 0.671, and the Significance is 0.000 (P<0.05), which meets the requirements of factor analysis (Table 6.1). Table 6.2 shows the rotated component matrix after factor analysis for the 12 perception features.

Table 6.1 Results of the Kaiser-Meyer-Olkin (KMO) test Bartlett's test of Sphericity.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.671
Bartlett's Test of Sphericity	Approx. Chi-Square	278.412
	df	66
	Sig.	.000

Table 6.2 Rotated Component Matrix.

	Component		
	1	2	3
Beautiful/Ugly	.884		
Meaningful/Meaningless	.884		
Artistic/Inartistic	.870		
Unique/Common	.864		
Easy to identify/Uneasy to identify	.808		
Continuous space/Interrupted space	.690		
Diversity/Monotonous	.663		
Natural/Artificial		.803	
Public/Private		.765	
Open/Enclosure		.738	
Rich green /Insufficient green		.634	
Modern/Historical			.866

Then we calculated each feature's weights using the entropy weight method and got the weights of each perception feature shown in Table 6.3.

Table 6.3 Weights of each feature.

Perception Features	weights
Public/Private	0.080
Natural/Artificial	0.079
Modern/Historical	0.080
Open/Enclosure	0.089
Diversity/Monotonous	0.069
Easy to identify/Uneasy to identify	0.059
Rich green/Insufficient green	0.101
Unique/Common	0.095
Beautiful/Ugly	0.071
Meaningful/Meaningless	0.089
Artistic/Inartistic	0.085
Continuous space/Interrupted space	0.102

Among the 12 perceptual features, we retained eight features with a weight>0.08 and removed the others because we had to find the main perception features by reducing the number of features. Then we got eight perception features. After merging with 17 main physical and image features obtained in Chapter 5 (Table 5.2), 25 main spatial features are shown in Table 6.4.

Table 6.4 Seventeen main physical, image, and perception features after reduction.

	Spatial features	Variable type
Physical features	Width between buildings (W_1) (m)	Continuous
	Width between boundary trees (W_2) (m)	Continuous
	Width between pavement's boundaries (W_3) (m)	Continuous
	Height of building's boundary (H_1) (m)	Continuous
	Height of boundary tree (H_2) (m)	Continuous
	W_1/H_1	Continuous
	W_2/H_2	Continuous
	Number of boundary layers	Ordinal
	Width of soft boundary	Continuous
	Number of space types	Ordinal
	Percentage of main space length	Continuous
	Number of material types	Ordinal
	Number of elements	Ordinal
Image features	Id_1wall-2building_26house	Continuous
	Id_3sky	Continuous
	Id_5tree_10grass_18plant	Continuous
	Id_7road_12sidewalk_53path_54stairs	Continuous
perception features	Public / Private	Continuous
	Modern / Historical	Continuous
	Open / Enclosure	Continuous
	Rich green / Insufficient green	Continuous
	Unique / Common	Continuous
	Meaningful / Meaningless	Continuous
	Artistic / Inartistic	Continuous
	Continuous space / Interrupted space	Continuous

6.3.2 Similarity and difference between the two countries

We used SPSS to analyze the significance of the difference in all the features between Japan and China.

First, the normal distribution test results of continuous variable features are shown that all the continuous variables are not normally distributed (Appendix H). So, we used the Mann-Whitney U-test to calculate the significance of the difference of all continuous variables between Japan and China. The Mann-Whitney test is the non-parametric equivalent of the independent samples t-test. This test will be used when the sample data is not normally distributed.

Third, we used the Chi-square test to calculate the significance of the difference of the ordinal variables (number of boundary layers, number of space types, number of material types, and number of elements) between Japan and China. Table 6.5 is a sample of the number of boundary layers for the Chi-square test.

Table 6.5 Sample of the data of the number of boundary layers for the Chi-square test.

.

	Number of boundary layers			
	1	2	3	Total
Japan	102	142	23	267
China	60	64	42	166
Total	162	206	65	433

Finally, we got the Asymptotic Significance (2-sided) of all features. To get the similarities and differences of each type of space, we calculated the median values of the main features of the ten spaces. We also calculated the median values of the main features of each type of public space.

Table 6.6 shows the value of the Asymptotic Significance (2-sided) between Japan and China and the median values of the main features of the ten spaces. The results show that there are no significant differences between the two countries in the following five features: Width between boundary trees (W2) (m), W2/H2, Id_ 1wall-

2building_ 26house, Modern / Historical, Artistic / Inartistic。 In other words, they are similar. The other 20 features are significantly different. Figures 6.1 and 6.2 compare the median values of physical, image, and perceptual features of ten public spaces in Japan and China.

Table 6.6 Values of the Asymptotic Significance (2-sided) between Japan and China and the medians of the main features in ten spaces.

	Spatial features	Asymptotic Significance (2-sided)	Median of features (Kanazawa and Nomi city)	Median of features (Dalian city)
Physical features	Width between buildings (W_1) (m)	0.000***	17.56	39.96
	Width between boundary trees (W_2) (m)	0.395	12.09	11.75
	Width between pavement's boundaries (W_3) (m)	0.000***	3.96	10.95
	Height of building's boundary (H_1) (m)	0.000***	7.46	12.50
	Height of boundary tree (H_2) (m)	0.000***	6.70	5.44
	W_1/H_1	0.000***	2.21	5.60
	W_2/H_2	0.195	1.82	2.61
	Number of boundary layers	0.006**	1.69	1.00
	Width of soft boundary	0.000***	11.41	4.34
	Number of space types	0.000***	2.25	1.75
	Percentage of main space length	0.000***	0.70	0.83
	Number of material types	0.000***	4.25	3.85
	Number of elements	0.000***	6.70	5.05
Image features	Id_1wall-2building_26house	0.392	0.19	0.14
	Id_3sky	0.168	0.13	0.11
	Id_5tree_10grass_18plant	0.018*	0.27	0.32
	Id_7road_12sidewalk_53path_54stairs	0.000***	0.18	0.27
perception features	Public / Private	0.000***	1.49	-0.88
	Modern / Historical	0.570	-0.64	-0.89
	Open / Enclosure	0.000***	1.10	-0.32
	Rich green / Insufficient green	0.000***	0.99	-0.45
	Unique / Common	0.000***	-0.47	-0.26
	Meaningful / Meaningless	0.000***	0.30	-0.45
	Artistic / Inartistic	0.757	0.13	0.04
	Continuous space / Interrupted space	0.000***	0.32	-1.19

*Note: * P<0.5; ** P<0.01; *** P<0.001.*

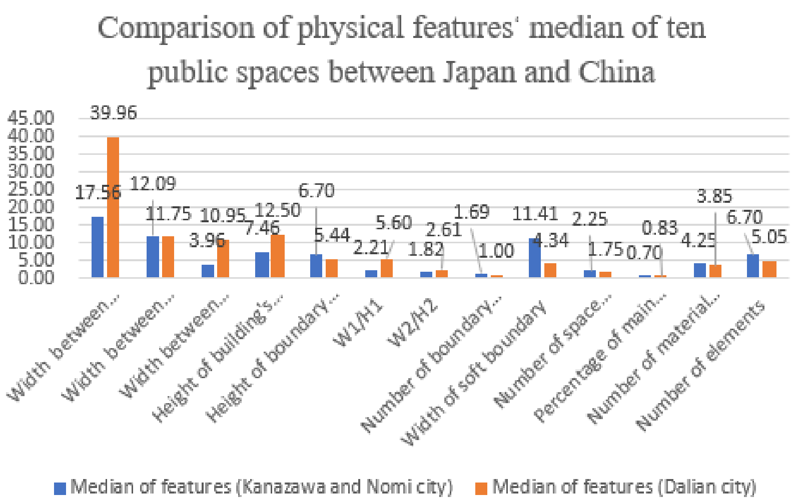


Figure 6.1 Comparison of physical features' medians of ten public spaces between Japan and China.

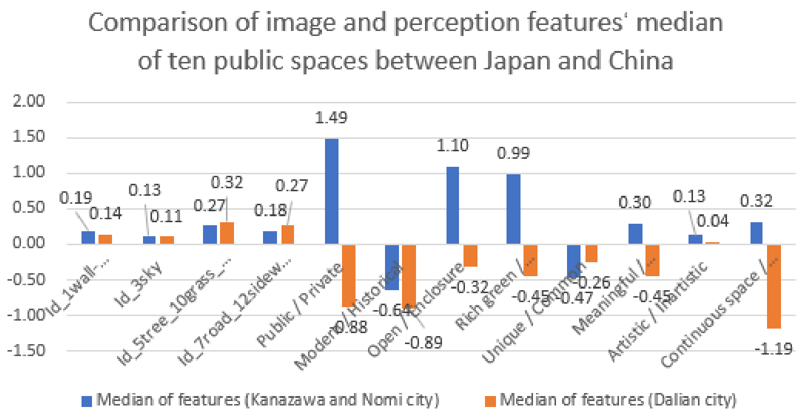


Figure 6.2 Comparison of image and perception features' medians of ten public spaces between Japan and China.

Table 6.7 shows the value of the Asymptotic Significance (2-sided) between Japan and China and the median values of the main features of the campus spaces. Figures 6.3 and 6.4 compare physical, image, and perception features' medians in the two campus spaces.

Table 6.7 Values of the Asymptotic Significance (2-sided) between Japan and China and the medians of the main features in two campus spaces.

	Spatial features	Asymptotic Significance (2-sided)	Median of features (Kanazawa and Nomi city)	Median of features(Dalian city)
Physical features	Width between buildings (W_1) (m)	0.078	36.00	51.50
	Width between boundary trees (W_2) (m)	0.330	8.00	13.50
	Width between pavement's boundaries (W_3) (m)	0.000***	4.00	6.00
	Height of building's boundary (H_1) (m)	0.001***	13.83	17.00
	Height of boundary tree (H_2) (m)	0.000***	2.33	7.00
	W_1/H_1	0.012*	2.18	3.03
	W_2/H_2	0.001***	2.42	1.81
	Number of boundary layers	0.000***	1.50	1.00
	Width of soft boundary	0.04*	2.00	5.50
	Number of space types	0.000***	3.00	2.00
	Percentage of main space length	0.143	0.57	0.62
	Number of material types	0.000***	5.00	3.00
	Number of elements	0.000***	7.00	5.00
Image features	Id_1wall-2building_26house	0.016*	0.15	0.05
	Id_3sky	0.000***	0.20	0.07
	Id_5tree_10grass_18plant	0.012*	0.23	0.38
	Id_7road_12sidewalk_53path_54stairs	0.006**	0.25	0.29
Perception features	Public / Private	0.000***	1.00	-1.27
	Modern / Historical	0.000***	0.89	-0.73
	Open / Enclosure	0.000***	0.67	-0.82
	Rich green / Insufficient green	0.000***	1.89	-1.55
	Unique / Common	0.000***	-0.78	0.55
	Meaningful / Meaningless	0.000***	-0.67	0.27
	Artistic / Inartistic	0.000***	-1.00	0.27
	Continuous space / Interrupted space	0.004**	0.22	-0.36

*Note: P<0.5; ** P<0.01; *** P<0.001.*

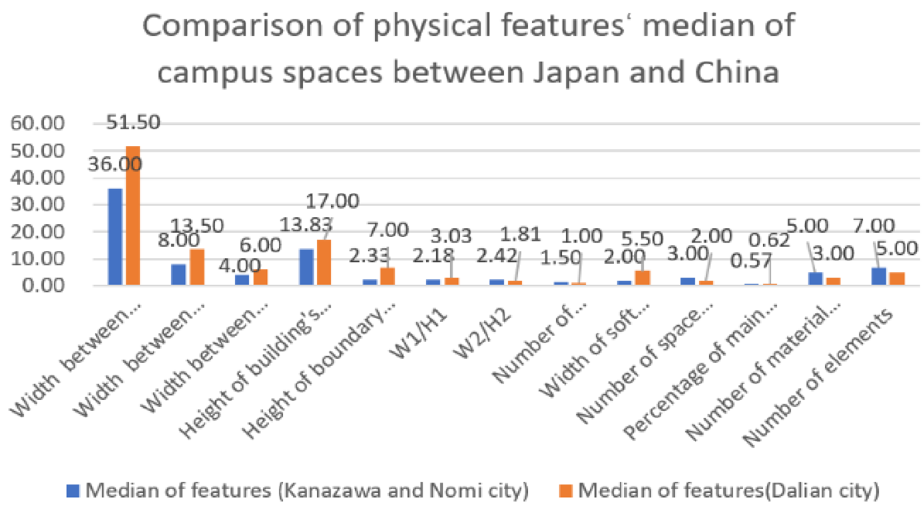


Figure 6.3 Comparison of physical features' medians of campus spaces.

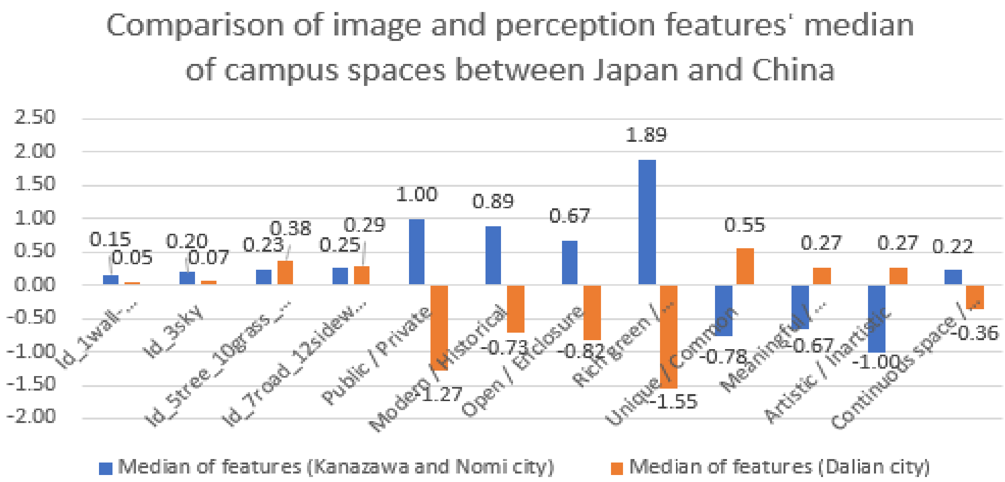


Figure 6.4 Comparison of image and perception features' median of campus spaces.

Table 6.8 shows the value of the Asymptotic Significance (2-sided) between Japan and China and the median values of the main features in the residential area spaces. Figures 6.5 and 6.6 compare physical, image, and perception features' medians in the two residential area spaces.

Table 6.8 Values of the Asymptotic Significance (2-sided) between Japan and China and the medians of the main features in two residential area spaces.

	Spatial features	Asymptotic Significance (2-sided)	Median of features (Kanazawa and Nomi city)	Median of features (Dalian city)
Physical features	Width between buildings (W_1) (m)	0.000***	12.50	24.00
	Width between boundary trees (W_2) (m)	0.032*	9.50	16.00
	Width between pavement's boundaries (W_3) (m)	0.001***	4.50	18.00
	Height of building's boundary (H_1) (m)	0.000***	7.00	15.00
	Height of boundary tree (H_2) (m)	0.000***	3.00	4.00
	W_1/H_1	0.000***	1.34	1.60
	W_2/H_2	0.032*	3.17	4.00
	Number of boundary layers	0.000***	1.50	1.25
	Width of soft boundary	0.270	3.00	2.50
	Number of space types	0.000***	3.00	1.50
	Percentage of main space length	0.000***	0.56	0.90
	Number of material types	0.000***	6.00	5.00
	Number of elements	0.000***	9.00	5.00
Image features	Id_1wall-2building_26house	0.000***	0.33	0.14
	Id_3sky	0.000***	0.16	0.09
	Id_5tree_10grass_18plant	0.000***	0.05	0.19
	Id_7road_12sidewalk_53path_54stairs	0.106	0.33	0.37
perception features	Public / Private	0.000***	2.56	-1.14
	Modern / Historical	0.061	-0.78	-0.73
	Open / Enclosure	0.000***	2.11	-0.95
	Rich green / Insufficient green	0.030*	-0.67	0.05
	Unique / Common	0.000***	-1.00	0.73
	Meaningful / Meaningless	0.000***	-1.56	0.36
	Artistic / Inartistic	0.000***	-2.11	1.23
	Continuous space / Interrupted space	0.002**	-2.00	-0.91

*Note: P<0.5; ** P<0.01; *** P<0.001.*

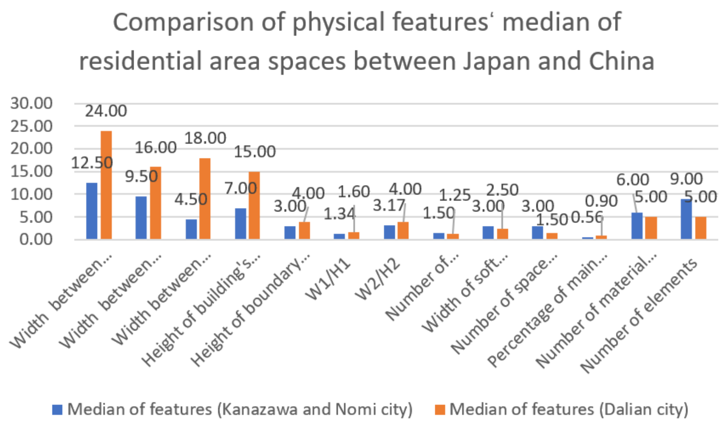


Figure 6.5 Comparison of physical features' medians of residential area spaces.

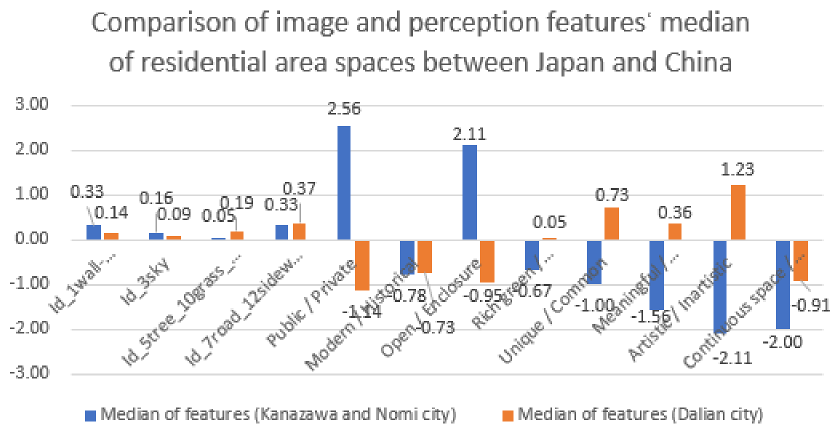


Figure 6.6 Comparison of image and perception features' median of residential area spaces.

Table 6.9 shows the value of the Asymptotic Significance (2-sided) between Japan and China and the median values of the main features in the park spaces. Figures 6.7 and 6.8 compare physical, image, and perception features' medians in the two park spaces.

Table 6.9 Values of the Asymptotic Significance (2-sided) between Japan and China and the medians of the main features in two park spaces.

	Spatial features	**Asymptotic Significance (2-sided)**	**Median of features (Kanazawa and Nomi city)**	**Median of features (Dalian city)**
Note:	Width between buildings (W_1) (m)	N/A	N/A	68.00
	Width between boundary trees (W_2) (m)	0.001***	8.33	35.00
	Width between pavement's boundaries (W_3) (m)	0.000***	5.00	26.00
	Height of building's boundary (H_1) (m)	N/A	N/A	12.00
	Height of boundary tree (H_2) (m)	0.000***	10.00	5.00
Physical features	W_1/H_1	N/A	N/A	5.67
	W_2/H_2	0.000***	1.25	7.00
	Number of boundary layers	0.856	2.25	2.00
	Width of soft boundary	0.000***	2.00	8.00
	Number of space types	0.010**	2.00	2.00
	Percentage of main space length	0.002**	0.76	0.76
	Number of material types	0.008**	4.00	4.00
	Number of elements	0.000***	7.00	6.00
Image features	Id_1wall-2building_26house	0.000***	0.00	0.10
	Id_3sky	0.510	0.10	0.10
	Id_5tree_10grass_18plant	0.000***	0.47	0.22
	Id_7road_12sidewalk_53path_54stairs	0.000***	0.13	0.28
perception features	Public / Private	0.000***	2.33	-0.55
	Modern / Historical	0.106	-1.33	-1.27
	Open / Enclosure	0.000***	2.11	0.09
	Rich green / Insufficient green	0.000***	2.56	-1.09
	Unique / Common	0.000***	1.67	-1.09
	Meaningful / Meaningless	0.000***	1.56	-1.36
	Artistic / Inartistic	0.000***	1.78	-0.91
	Continuous space / Interrupted space	0.000***	1.11	-1.27

** $P<0.5$; ** $P<0.01$; *** $P<0.001$.*

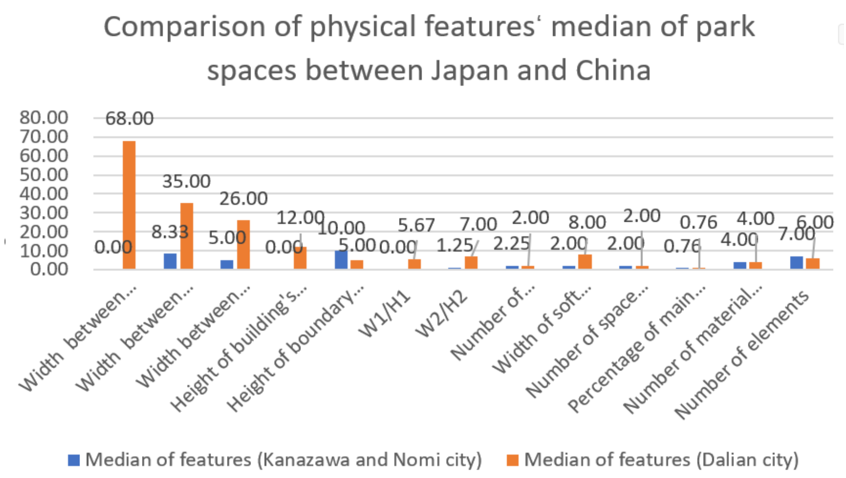


Figure 6.7 Comparison of physical features' medians of two park spaces.

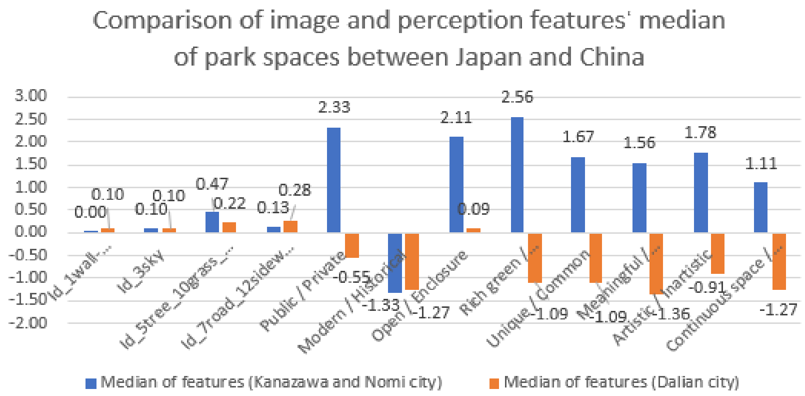


Figure 6.8 Comparison of image and perception features' median of two park spaces.

Table 6.10 shows the value of the Asymptotic Significance (2-sided) between Japan and China and the median values of the main features in the memorial spaces. Figures 6.9 and 6.10 compare physical, image, and perception features' medians in the two park spaces.

Table 6.10 Values of the Asymptotic Significance (2-sided) between Japan and China and the medians of the main features in two memorial spaces.

	Spatial features	Asymptotic Significance (2-sided)	Median of features (Kanazawa and Nomi city)	Median of features (Dalian city)
	Width between buildings (W_1) (m)	0.000***	5.00	N/A
	Width between boundary trees (W_2) (m)	0.356	1.83	3.00
Note:	Width between pavement's boundaries (W_3) (m)	0.000***	1.33	2.00
	Height of building's boundary (H_1) (m)	0.323	3.00	N/A
	Height of boundary tree (H_2) (m)	0.019*	6.00	6.00
Physical features	W_1/H_1	0.000***	2.50	N/A
	W_2/H_2	0.921	0.31	0.50
	Number of boundary layers	0.000***	2.00	3.00
	Width of soft boundary	0.000***	1.50	3.00
	Number of space types	0.000***	2.00	1.00
	Percentage of main space length	0.000***	0.63	1.00
	Number of material types	0.000***	3.00	4.00
	Number of elements	0.005**	6.00	6.00
Image features	Id_1wall-2building_26house	0.000***	0.13	0.02
	Id_3sky	0.000***	0.01	0.06
	Id_5tree_10grass_18plant	0.000***	0.29	0.55
	Id_7road_12sidewalk_53path_54stairs	0.000***	0.01	0.17
perception features	Public / Private	0.723	-0.56	-0.82
	Modern / Historical	0.203	2.78	-0.55
	Open / Enclosure	0.001***	-1.22	0.36
	Rich green / Insufficient green	0.000***	0.33	-1.45
	Unique / Common	0.000***	-2.44	-1.09
	Meaningful / Meaningless	0.000***	-2.67	-1.00
	Artistic / Inartistic	0.000***	-2.89	-0.18
	Continuous space / Interrupted space	0.000***	1.67	-1.18

** P<0.5; ** P<0.01; *** P<0.001.*

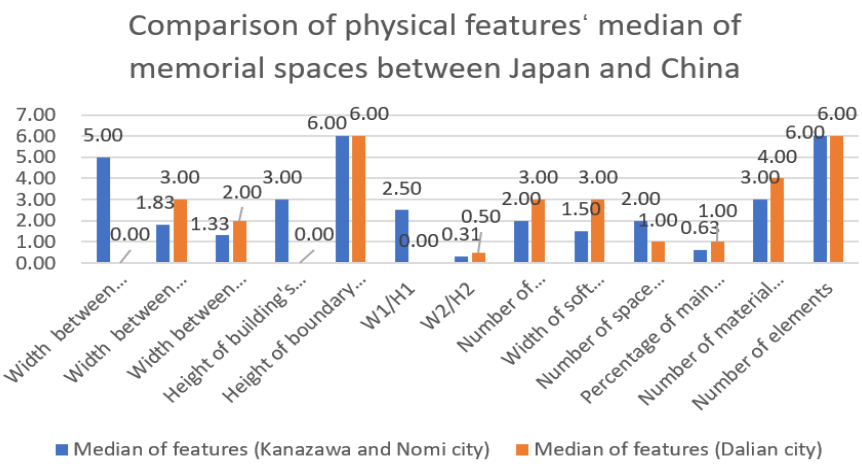


Figure 6.9 Comparison of physical features' medians of two memorial spaces.

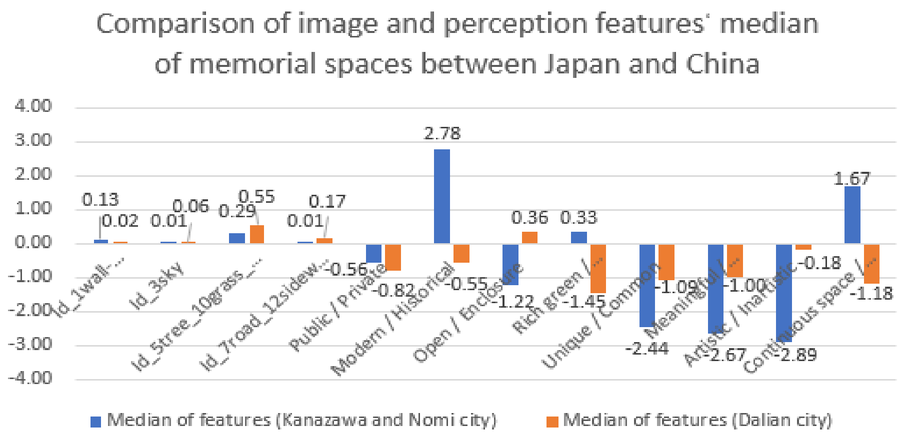


Figure 6.10 Comparison of image and perception features' median of two memorial spaces.

Table 6.11 shows the value of the Asymptotic Significance (2-sided) between Japan and China and the median values of the main features in the pedestrian street spaces. Figures 6.11 and 6.12 compare physical, image, and perception features' medians in the two pedestrian street spaces.

Table 6.11 Values of the Asymptotic Significance (2-sided) between Japan and China and the medians of the main features in two pedestrian street spaces.

	Spatial features	Asymptotic Significance (2-sided)	Median of features (Kanazawa and Nomi city)	Median of features(Dalian city)
Physical features	Width between buildings (W_1) (m)	0.000***	4.33	15.00
	Width between boundary trees (W_2) (m)	0.000***	10.00	22.00
	Width between pavement's boundaries (W_3) (m)	0.000***	2.67	8.00
	Height of building's boundary (H_1) (m)	0.000***	7.00	11.00
	Height of boundary tree (H_2) (m)	0.000***	9.67	2.50
	W_1/H_1	0.005**	0.62	1.71
	W_2/H_2	0.000***	1.03	8.80
	Number of boundary layers	0.038*	1.00	1.00
	Width of soft boundary	0.389	1.00	0.00
	Number of space types	0.000***	2.00	1.00
	Percentage of main space length	0.022*	0.74	0.95
	Number of material types	0.000***	4.00	5.00
	Number of elements	0.000***	7.00	6.00
Image features	Id_1wall-2building_26house	0.002**	0.50	0.38
	Id_3sky	0.001***	0.06	0.13
	Id_5tree_10grass_18plant	0.159	0.01	0.02
	Id_7road_12sidewalk_53path_54stairs	0.000***	0.16	0.33
perception features	Public / Private	0.000***	1.44	-1.45
	Modern / Historical	0.160	-1.89	-2.00
	Open / Enclosure	0.000***	0.56	-0.36
	Rich green / Insufficient green	0.000***	-0.44	1.73
	Unique / Common	0.578	-0.33	0.09
	Meaningful / Meaningless	0.578	-0.78	0.27
	Artistic / Inartistic	0.004**	-0.89	-0.45
	Continuous space / Interrupted space	0.000***	0.11	-1.27

*Note: * P<0.5; ** P<0.01; *** P<0.001.*

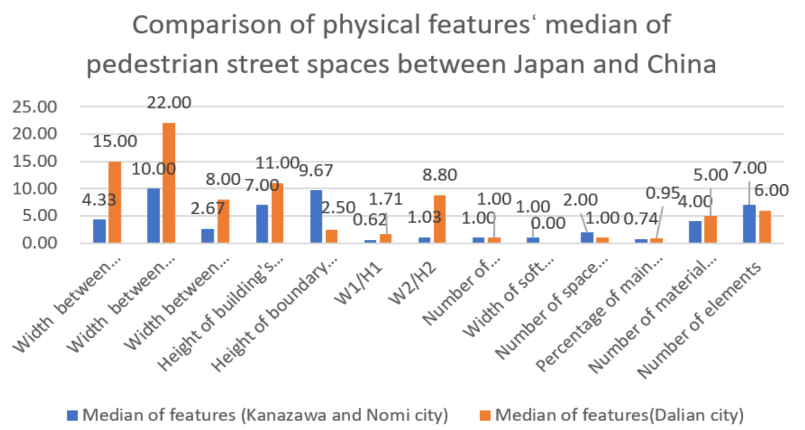


Figure 6.11 Comparison of physical features' medians of two pedestrian street spaces.

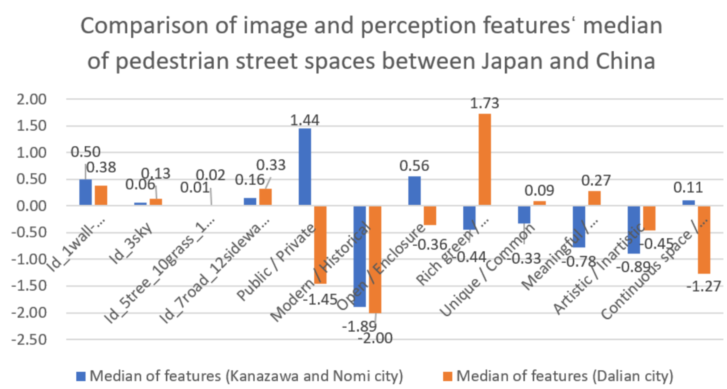


Figure 6.12 Comparison of image and perception features' median of two pedestrian street spaces.

6.4 Discussion

1) There are more differences than similarities between the two countries, and among the 19 features with significant differences, the median difference of 8 features is large.

First, among the 25 features of 10 spaces, only six features are not different between Japan and China; that is, they are similar, and the remaining 19 features are different, which shows that there are many differences in the features of public spaces between Japan and China. The differences are mainly manifested in the W_1/H_1, the width of the road, the width of the soft border, the openness of the space, and the proportion of vegetation. However, the proportion of the sky, the modernity or tradition of the space, and the artistry are similar. These results indicate that the spatial scales enclosed by the soft boundaries between Japan and China are similar (W2/H2 in Japan and China are 1.82 and 2.61, respectively), but the width between buildings (W1) in Chinese is larger than that in Japan (17.56m and 39.96m in Japan and China, respectively). The Id_5tree_10grass_18plant is similar (0.27 and 0.32 in Japan and China, respectively), but the number of boundary layers is different (1.69 and 1.00 in Japan and China, respectively). The width of the soft boundary is about three times that of China (11.41m and 4.34m in Japan and China, respectively). And the W1/H1 of the space is about twice that of Japan (2.21 and 5.60 in Japan and China, respectively), the boundary design focuses on the combination of soft and hard boundaries.

2) Among the five types of public spaces, the differences among campus, residential areas, and memorial sites are more than between parks and pedestrian streets.

a. Comparing the features of campus space, we found that the median values of uniqueness and artistry in Japanese campuses are relatively high (the uniqueness of Japan and China are -0.778 and 0.545, and the artistry is -1.000 and 0.273, respectively). The proportion of the sky in Japan is higher than that of China (0.152 and 0.050, 0.196 and 0.066 respectively in Japan and China), and the richness of the Chinese campus in the soft boundary may be related to the history of the two campuses (JAIST was built in 1990, while DLPU was built in 1958).

b. Through comparing the features of residential areas, we found that the two countries have similar soft border widths (3.000 and 2.500 in Japan and China, respectively), but the proportion of trees in residential areas in Japan is less than that in China (0.047 and 2.500 in Japan and China, respectively). 0.194). Regarding perceptual features, the greenness of residential areas in Japan is higher than that in China (-0.667 and 0.045 in Japan and China, respectively). This may be due to China's relatively simple layer and boundary treatment of plants.

c. Comparing the park spaces between Japan and China, we found that the W1/H1 of the two are very different (1.250 and 7.000 for Japan and China, respectively). The width of the soft boundary between Japan and China is 2.000 and 8.000, respectively, which shows that the boundary transition of Chinese parks is better. However, the proportion of greenery in Japanese parks is relatively high (0.472 and 0.224 in Japan and China, respectively), which may be related to the styles of the two park samples; and Japanese parks are traditional Parks, Chinese parks are modern parks.

d. In the monumental space, the W2/H2 (0.310 and 0.500), privacy (-0.556 and -0.818), and modernity (2.778 and -0.545) in Japan and China are similar. However, there are large differences in image features. In addition, the median values of uniqueness, significance, and artistry of Japanese monumental spaces are all high. This shows that the two countries have different ways of expressing commemoration. China pays attention to naturalness, while Japan pays more attention to the expression of meaning.

e. The comparison of the features of the pedestrian street space in Japan and China shows that the W1/H1 (4.333 and 15.000), the proportion of the sky (0.063 and 0.133), and the artistry (-0.889 and -0.455) are pretty different, but the unique/common (-0.333 and 0.091) and meaningful/meaningless (-0.778 and 0.273) were similar. These results indicate that the two are different in the spatial scale of pedestrian streets but similar in spatial perception.

After comparing each type of space, we found that the five types of space between Japan and China have more similarities in Width between boundary trees (W2) and modern/traditional. Moreover, there are more similarities in the spatial features of parks and pedestrian streets (5 similar features), while there are more differences in the other three types of spaces. (22 different features).

3) Compared with related studies, this study found the main features that affect the difference between Japanese and Chinese public spaces and the degree of the difference between the median values.

Different from relevant studies (Chen, 1991; Zhao et al., 2003; Mossman, 2009; Han, 2022), we compared Japanese and Chinese public spaces from multiple types of public spaces and found more differences than similarities among the 25 features between Japan and China. The main features that affect the difference are spatial scale, boundary, public/private, and continuity of space. The feature quantification comparison in this study is a continuation and supplement of previous studies.

6.5 Summary

In this chapter, we explored the similarities and differences in the features between Japan and China through the comparative analysis of the data on the public spaces' physical, image, and perceptual features. The results show that the differences in the features of public spaces between Japan and China are more than similar. The differences are mainly manifested in the W1/H1, the width of the road, the width of the soft border, the openness, and the artistry of the space. Analyzing the similarities and differences in the two countries' public spaces gives us a more comprehensive understanding of the public spaces in Japan and China.

Chapter 7
Conclusion, Implication, and Limitations

7.1 Conclusion

The primary purpose of this dissertation is to explore the relationship between the features of public space and user emotions. We proposed three specific sub-goals: sub-goal 1 is to build a multi-type public space emotion classification model oriented to design practice; sub-goal 2 is to extract quantitative features of the space with high emotion-eliciting quality; sub-goal 3 is to compare the similarity and differences between Japan and China. The following are the conclusions on the three goals:

7.1.1 Conclusion for the sub-goals

1) For sub-goal 1: this study collected physiological signals (EDA, ECG, EMG) and self-assessment manikin (SAM) data from 20 participants in 10 public spaces of 5 categories in three cities in Japan and China. Then we use six classifiers, LR, DT C5.0, ANN, DT C5.0 (boosting), RF (bagging), and ANN (boosting), to build the models. Meanwhile, we introduced the synthetic minority oversampling technique (SMOTE) to solve the problem of sample imbalance. Finally, we built space emotion classification models suitable for multi-type spaces. From the results, we could find that the highest accuracies of the binary and ternary classification models were 94.29% and 91.07%, respectively. After external validation, the highest accuracies were 80.90% and 65.30%, respectively, which satisfied the preliminary requirements for evaluating the quality of existing urban spaces. However, the quinary classification model could not meet the primary needs. Second, the average accuracy of ensemble learning was 7.59% higher than that of single classifiers. Third, the application of SMOTE solved the problem of overfitting caused by excessive reliance on a minimal number of samples and the problem of poor accuracy and adaptability of the model in practical application.

2) For sub-goal 2: this study introduced the fully convolutional neural network (FCN) to do semantic segmentation on spatial images, combined with the physical features of space extracted in real space to form feature variable datasets. Then five clusters were

obtained by using an unsupervised learning two-step clustering analysis. By comparing the distribution ranges of these spatial features, we finally got the range of values associated with the high and low valence. The results show that nine features are the main features that affect emotional valence; in other words, these nine main features might work together to distinguish whether the space is a positive space or a negative space, and the other features play an auxiliary role. The value range of these spatial features might support new and reconstruction projects in urban design.

3) For sub-goal 3: through comparing the physical, image, and perceptual features of Japanese and Chinese public spaces, the results show that the differences in the features of public spaces between Japan and China are more than similar. The differences are mainly manifested in the W1/H1, the width of the road, the width of the soft border, the openness, and the artistry of the space. Analyzing the similarities and differences in the two countries' public spaces gives us a more comprehensive understanding of the public spaces in Japan and China.

7.1.2 Conclusion for the main goal

First, from the above description of the three sub-goals, we not only make it clear that there is an association between the features of public space and the emotional response of users but also that different types of public spaces will have similar results for users.

Second, the physiological signals realized the quantifiable expression of participants' emotions, and physical and image features of public spaces realized the quantifiable expression of spatial features; thus, we found data-based evidence for the association between emotion and spatial features. This association includes the classification model and the main spatial features of the positive and negative spaces.

Third, the research results of sub-goal 1 show that we could recognize others' emotional reactions to public space through the emotional classification models. We got better classification accuracy with the help of multiple physiological signals and ensemble classifiers. Combining the research results of sub-goals 1 and 2, we found the corresponding relationship between spatial features and the user's emotions and features. The corresponding relationship might help us understand their relationship and support urban design decisions.

7.2 Research implications

7.2.1 Application process for urban design

The quality evaluation and feature extraction of urban public space based on data evidence are issues of built space in cities. In the project practice, it is difficult for the clients and designers to agree on the built space's quality evaluation because both parties' evaluation is based on the personal evaluation. Therefore, our research aims to obtain the spatial quality evaluation model based on data evidence and the common features of high-quality spaces（Figure 7.1 ）.

The results of the study of sub-goal 1 in this dissertation support evaluating the emotion-eliciting quality and management decisions of public space, including the feasibility evaluation before construction and the post-occupancy evaluation after construction.

The results of sub-goal 2 support ideation decisions in the overall scheme and detailed design phases, including evaluating existing spatial features and comparing features before and after renewal.

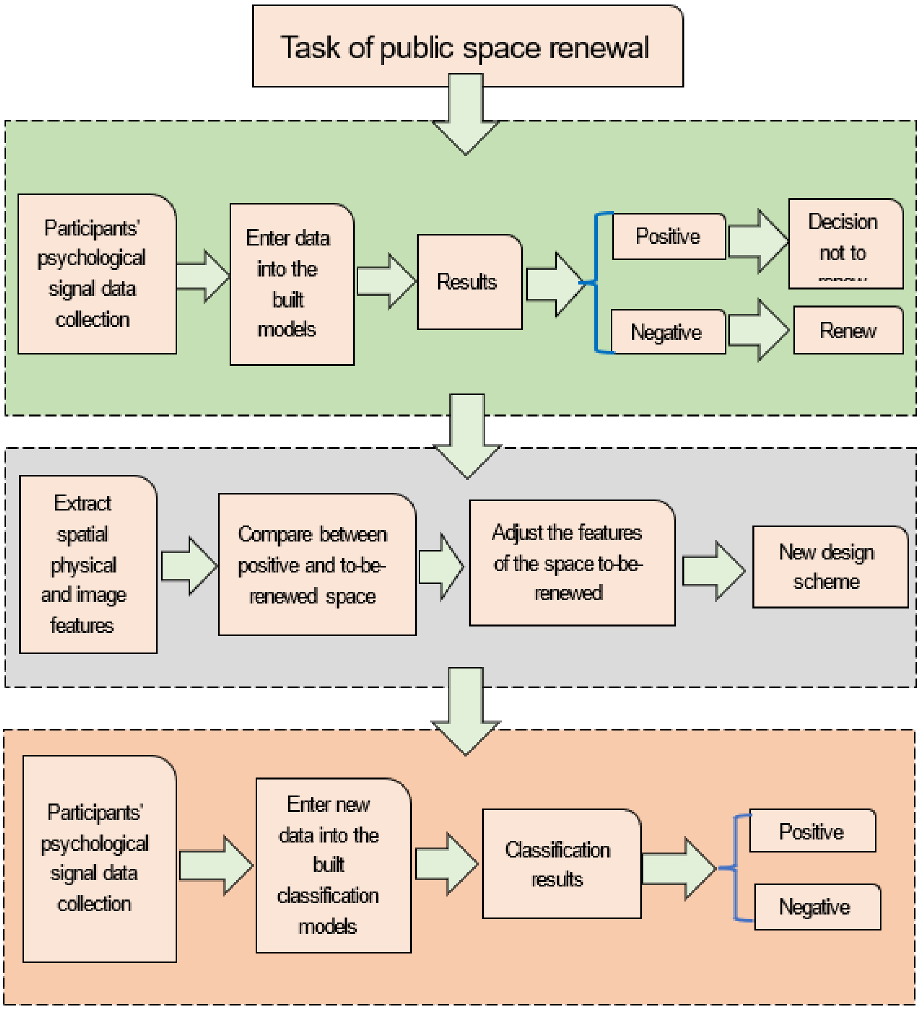


Figure 7.1 The application process of studies 1 and 2 to renew the public spaces.

1) For the built public spaces, the study of sub-goal 1 provides a method to diagnose emotion-eliciting quality. Unlike previous studies that only support single-type spaces, this method is suitable for multi-type spaces.

The debate on whether the evaluation of public space should pay more attention to the opinions of experts (Craig et al., 2002; Mehta, 2013; Zhang et al., 2018; Tang and Long, 2019), users (Li et al., 2016; Ernawati et al., 2016; Fathullah and Willis, 2018; Bivina et al., 2018) or both (Mehta, 2013; Steinmetz-Wood, 2019) has been inconclusive. These evaluation methods are not only greatly affected by the personal factors of the participants but also challenging to determine the weight of the features. The evaluation method proposed in the study of sub-goal 1 uses the physiological model as the evaluation index, which reduces personal factors' impact and avoids the weight problem.

The results of this study could be suitable for the public space to be renewed. Generally, such projects need to solve two problems: determining which public spaces need to be renewed and how to redesign the spaces. Figure 7.1 shows that combining the study of sub-goals 1 and 2 helps solve these two problems. The spatial emotion recognition model established in the study of sub-goal 1 solves the first problem and then applies the results of the study of sub-goal 2 to determine which spatial features need to be changed to produce a positive public space. After rebuilding the space according to the new design scheme, we can apply the method of the study of sub-goal 1 again for post-occupancy evaluation. If the evaluation result is negative, we go back to the design stage to adjust the scheme until we get a positive result.

It may be feasible to use this process to evaluate the quality of spatial emotional stimulation in public spaces. However, in the specific operation, it needs a professional operation, including data collection, analysis, and obtaining results. Finally, the professional organization will submit the report to the project management organization as part of the feasibility report.

2) The basic features of positive space will be used to judge which features need to be changed, and then we will design and renew the public space by combining them with basic and special features.

The features of public spaces are diverse, and it seems infeasible to try to find all the features that affect emotion-eliciting quality. Therefore, we try to reduce the dimensionality of the spatial features and find the main features that affect the spatial quality, that is, the basic features (as opposed to the basic features are special features). High- and low-quality public spaces have different basic features (Table 7.1), while special features make the space individual and unique. Therefore, in the project practice, after determining the design aim, we divide the space and then select the basic features of each space according to the needs of the space sequence to form the basic quality and atmosphere of the space. Finally, we add special features to form the style and uniqueness of the space (Figure 7.2).

Table 7.1 Diagnostic table of positive public space features.

	Features	Positive public space		Results
		Range	Median	
Main fetures	Id_5tree_10grass_18plant	11.00%-49.00%	32.00%	☐ Yes ☐ No
	W2/H2	0.50-7.00	3.26	☐ Yes ☐ No
	Width between boundary trees (W2) (m)	4.00-22.00	22.00	☐ Yes ☐ No
	Width of soft boundary (m)	3.00-11.60	5.50	☐ Yes ☐ No
	W1/H1	0.95-3.03	1.91	☐ Yes ☐ No
	Percentage of main space length	62.00%-84.00%	63.00%	☐ Yes ☐ No
	Number of boundary layers	2.00-3.00	2.50	☐ Yes ☐ No
	Width between pavement boundaries (W3) (m)	3.33-6.00	3.33	☐ Yes ☐ No
	Width between buildings (W1) (m)	5.00-51.50	6.67	☐ Yes ☐ No
Secondary fetures	Id_1wall-2building_26house	1.00%-33.00%	9.00%	☐ Yes ☐ No
	Number of material types	3.00-4.00	3.00	☐ Yes ☐ No
	Id_7road_12sidewalk_53 path_54stairs	11.00%-30.00%	22.00%	☐ Yes ☐ No
	Height of buildings boundary (H1) (m)	2.00-12.00	7.00	☐ Yes ☐ No
	Number of elements	5.00-7.00	6.00	☐ Yes ☐ No
	Height of boundary tree (H2) (m)	5.00-8.00	6.00	☐ Yes ☐ No
	Number of space types	1.00-3.00	2.00	☐ Yes ☐ No
	Id_3sky	2.00%-17.00%	9.00%	☐ Yes ☐ No

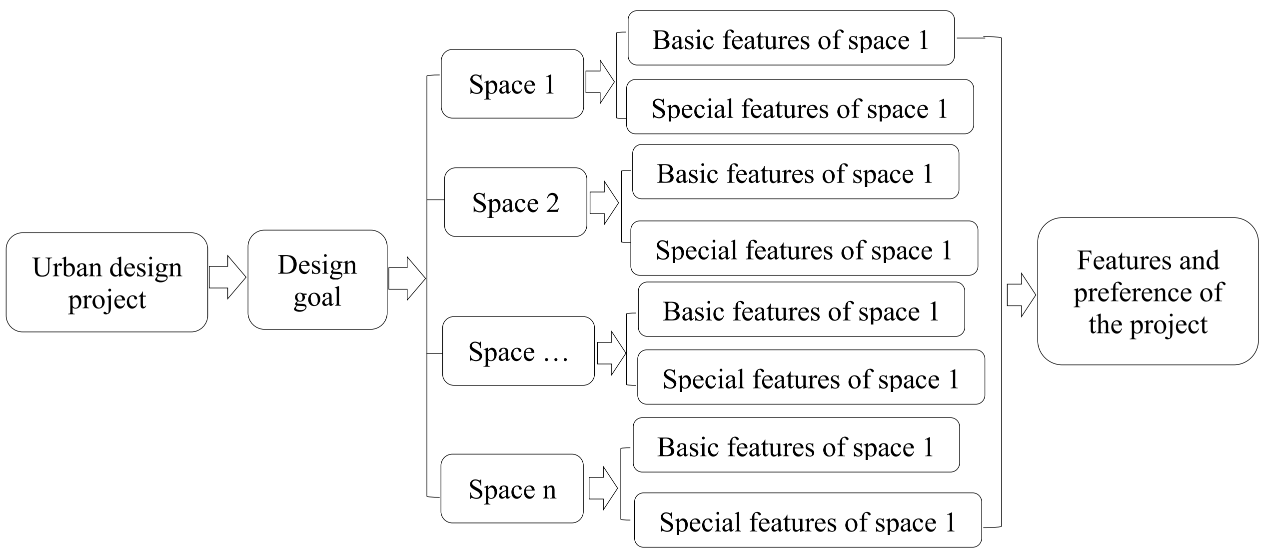


Figure 7.2 Application process of basic features of high- and low-quality public space in actual projects.

7.2.2 Contribution to knowledge science

Knowledge science is a problem-oriented interdisciplinary research field. One of its aims is to break through the boundaries of disciplines and comprehensively apply research methods and technologies of humanities, social sciences, cognitive sciences, and information sciences to solve problems.

1) Using the methods of urban science, psychology, and information science methods, this study built a sample of the emotion-eliciting quality model of multiple public spaces. The model sample used emotional and physiological signal data to judge the quality of space, supporting the renewal decision-making of urban public spaces.

2) Compared with the previous urban space diagnosis methods that used professional evaluation indicators, the method based on the user’s emotions is an integral evaluation of the relationship between people and the environment. This method is easily understood or accepted by people from different backgrounds, including managers and the public; It also facilitates knowledge transfer and exchange among designers, managers, and users.

3) It provides evidence based on data for urban space renewal and design. The previous design was based on the analysis of the current situation and the designer's judgment. This method depends on the ability of the designer's individual or team, which might be unreliable. The quantitative diagnostic table based on the main features of the functional space of data proposed in this study will probably avoid such unreliability. This diagnostic table cannot cover all aspects of space problems, but it guarantees the basic quality requirements of urban public space from the feature level.

4) We compared the features between Japan and China through the two-dimensional images, the three-dimensional spaces, and the psychological perception and obtained the quantitative comparison results. This method is different from related research and could be used to solve the other problem about urban space, for example, comparing the spaces between new and old or spaces between different regions.

7.3 Limitations and future works

1) Limitations

This study built an emotion-eliciting quality evaluation model for multi-type urban public spaces. However, the proposed model had limitations in the following five aspects.

First, the previous research was to collect data in one space. Our research was to collect data in five types of spaces, which expanded the scope of the application of the model. So, it is necessary to collect data from more types of spaces, such as waterfront spaces, squares, and urban streets, to establish a model with more application value;

Second, the participants are Chinese students in JAIST (including five master's students and four doctoral students) and 11 master's students at Dalian Polytechnic University. Therefore, the background and age of the participants are relatively simple, so it is necessary to increase the diversity of participants.

Third, the method collected the photos of spatial routes in this research was that the camera shooting direction is parallel to the route, so this method could only capture a part of the objects in the spaces, but users will turn their heads at will and look at different angles. So, this is the limitation of our research. In future research, we will try to solve this problem with a 360-degree camera, and the new data collection method maybe improves our research.

Fourth, the data of emotion-eliciting quality assessment could not reflect the comprehensive features of the public space because it was based on personal experience. Therefore, commercial and spatial behavior data will be added to the evaluation model.

Fifth, this study built several emotion-eliciting quality evaluation models for multi-type urban public spaces. However, human emotions include short-term and long-term effects. Users who enter a public space for the first time rely primarily on their physical senses to perceive it. After long-term use, factors such as space function, public social interaction, and place attachment become the main factors affecting evaluation. Thus, it is necessary to analyze further the long-term emotions evoked by a space to obtain a more comprehensive assessment of its affective quality.

2) Future works

Given the above limitations, we will continue to improve our research methods in the future to make progress in the research of the relationship between public space and user emotions.

First, more spatial data and participants with different backgrounds should be added to improve the reliability and applicability of research results.

Second, we will test the established model in practice to verify its effectiveness of the model, including tests in the built space and new space. Furthermore, we will cooperate with urban design companies to help designers try to apply the space quality diagnosis table in the design decision-making process to verify the validity of the value range and median of positive spatial features.

Third, we will try to reduce the impact of devices on participants in the data collection stage. We may obtain the correlation between spatial features and emotional arousal by comparing various spaces.

Fourth, we will try to combine static and dynamic collection to remove the data noise caused by walking. At the same time, we will increase the types of space, including streets, street green spaces, sports parks, suburban parks, greenways, and other project types, to expand the model's adaptability.

Fifth, although the binary classification model can be used to evaluate multiple types of public spaces, the results of the ternary and quinary classifications were poor. So, we will continue to improve the ability of the ternary and quinary classifications and try to test the effectiveness of the model in practice.

In sum, we will attempt to study the effects of long-term emotions, spatial function, and neighborhood interaction on evaluating spatial affective quality. Through multimodal signal extraction and new machine learning technologies, we will continuously improve the performance of the spatial quality evaluation model and provide technical support for the construction of intelligent cities.

Reference

Al Machot, F., Elmachot, A., Ali, M., Al Machot, E. and Kyamakya, K. (2019). A Deep-Learning Model for Subject-Independent Human Emotion Recog-nition Using Electrodermal Activity Sensors. Sensors, 19(7), p.1659. doi:10.3390/s19071659.

Al-barrak, L., Kanjo, E. and Younis, E.M.G. (2017). NeuroPlace: Categorizing urban places according to mental states. PLOS ONE, 12(9), p.e0183890. doi:10.1371/journal.pone.0183890.

Alberdi, A., Aztiria, A. and Basarab, A. (2016). Towards an automatic early s-tress recognition system for office environments based on multimodal measu-rements: A review. Journal of Biomedical Informatics, 59, pp.49–75. doi:10.1016/j.jbi.2015.11.007.

Ali, M., Machot, F., Mosa, A., Jdeed, M., Machot, E. and Kyamakya, K. (2018). A Globally Generalized Emotion Recognition System Involving Different Physiological Signals. Sensors, 18(6), p.1905. doi:10.3390/s18061905.

Alzoubi, O., Hussain, Md.S., D'Mello, S. and Calvo, R.A. (2011). Affective Modeling from Multichannel Physiology: Analysis of Day Differences. Affec-tive Computing and Intelligent Interaction, pp.4–13. doi:10.1007/978-3-642-24600-5_4.

Antoine Couturier. 2017. EMG Toolbar (ver. 5.30). [Software]. [Accessed 2020.04.22]. Retrieved from https://www.originlab.com/fileExchange/details.aspx?fid=420.

Appelhans, B.M. and Luecken, L.J. (2006). Heart Rate Variability as an Index of Regulated Emotional Responding. Review of General Psychology, 10(3),

pp.229–240. doi:10.1037/1089-2680.10.3.229.

Arnold, M. B. (1950). An excitatory theory of emotion. In M. L. Reymert,

Feelings and emotions: The Mooseheart Symposium in cooperation with the University of Chicago (pp. 11-33). New York: McGraw-Hill.

Banda N, Engelbrecht A and Robinson P. (2015) Feature reduction for dimensi-onal emotion recognition in human robot interaction. In: 2015 IEEE Sympo-sium Series on Computational Intelligence. IEEE; pp. 803–810. https://doi.org/10.1109/SSCI.2015.119

Bergner, B., Exner, J., Memmel, M., Raslan, R., Talal, M., Taha, D., and Zeile,P. (2013) Human sensory assessment methods in urban planning–A case stu- dy in Alexandria. In PLANNING TIMES–You Better Keep Planning or YouGet in Deep Water, for the Cities They are A-Changin, Proceedings of the 18th International Conference on Urban Planning, Regional Development, and Information Society; 20-23 May, Rome, Italy, pp. 407-417.

BIOPAC Systems, Inc.. (2018). Acqknowledge (ver. 4.2). [Software]. [Accessed

2020.03.12]. Retrieved From https://www.biopac.com/acqknowledge-4-2-released.

Birenboim, A., Dijst, M., Scheepers, F.E., Poelman, M.P. and Helbich, M. (2019). Wearables and Location Tracking Technologies for Mental-State Sensing in Outdoor Environments. The Professional Geographer, 71(3), pp.449–461. doi:10.1080/00330124.2018.1547978.

Bivina, G., Parida, P., Advani, M., and Parida, M. (2018) Pedestrian LOS model for evaluating and improving sidewalks from various land uses. European Transport-Trasporti Europei 67(2): pp.1-18.

Blair, K.S., Smith, B.W., Mitchell, D.G.V., Morton, J., Vythilingam, M., Pessoa,L., Fridberg, D., Zametkin, A., Nelson, E.E., Drevets, W.C., Pine, D.S.,

Martin, A. and Blair, R.J.R. (2007). Modulation of emotion by cognition and cognition by emotion. NeuroImage, 35(1), pp.430–440. doi:10.1016/j.neur-oimage.2006.11.048.

Bradley, M.M. and Lang, P.J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry, 25(1), pp.49–59. doi:10.1016/0005-7916(94)90063-9.

Brosch, T., Pourtois, G. and Sander, D. (2010). The perception and categorisat-ion of emotional stimuli: A review. Cognition & Emotion, 24(3), pp.377–400.doi:10.1080/02699930902975754.

Cannon, W.B. (1987). The James-Lange Theory of Emotions: A Critical Exami-nation and an Alternative Theory. The American Journal of Psychology, 100(3/4), p.567. doi:10.2307/1422695.

Carmona, M., Heath, T., Tiesdell, S., and Oc, T. (2010) Public places, urban spaces: the dimensions of urban design. London: Routledge.

Carr, S., Francis, M., Rivlin, L.G. and Stone, A.M. (1992). Public space. Cam-bridge University Press.

Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002). SMOTE:Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligen-ce Research, 16(16), pp.321–357. doi:10.1613/jair.953.

Chen, D., Burley, J.B., Machemer, T. and Schutzki, R. (2021b). Ordination of Selected Traditional Japanese Gardens, Traditional Chinese Gardens, and Modern Chinese Gardens. International Journal of Culture and History, 8(1), p.14. doi:10.5296/ijch.v8i1.18250.

Chen, J., Hu, B., Wang, Y., Moore, P., Dai, Y., Feng, L. and Ding, Z. (2017). Subject-independent emotion recognition based on physiological signals: a three-stage decision method. BMC Medical Informatics and Decision Making, 17(S3). doi:10.1186/s12911-017-0562-x.

Chen, S., Jiang, K., Hu, H., Kuang, H., Yang, J., Luo, J., Chen, X. and Li, Y. (2021a).

Emotion Recognition Based on Skin Potential Signals with a Portable Wireless Device. Sensors, 21(3), p.1018. doi:10.3390/s21031018.

Cho, M.E. and Kim, M.J. (2017). Measurement of User Emotion and Experien-ce in Interaction with Space. Journal of Asian Architecture and Building En-gineering, 16(1), pp.99–106. doi:10.3130/jaabe.16.99.

Christelle G, Fabrice PB, Aure´lie C, Sylvie Charbonnier, Ste´phane Bonnet and Audrey Vidal. (2015) Features relevance analysis for emotion classificat-ion with physiological sensors. In: Proceedings of the 2nd International Conference on Physiological Computing Systems. Scitepress—Science and Technology Publications; pp. 17–25. https://doi.org/10.5220/0005238600170025

Collins, G.S., de Groot, J.A., Dutton, S., Omar, O., Shanyinde, M., Tajar, A., Voysey, M., Wharton, R., Yu, L.-M., Moons, K.G. and Altman, D.G. (2014). External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. BMC Medical Research Methodology, 14(1). doi:10.1186/1471-2288-14-40.

Consonni, V., Ballabio, D. and Todeschini, R. (2010). Evaluation of model predictive ability by external validation techniques. Journal of chemometrics, 24(3-4), pp.194-201.

Craig, C.L., Brownson, R.C., Cragg, S.E. and Dunn, A.L. (2002). Exploring the effect of the environment on physical activity. American Journal of Preventive Medicine, 23(2), pp.36–43. doi:10.1016/s0749-3797(02)00472-5.

Davidson, J. and Milligan, C. (2004). Embodying emotion sensing space: introducing emotional geographies. Social & Cultural Geography, 5(4), pp.523–532. doi:10.1080/1464936042000317677.

Doherty, S.T., Lemieux, C.J. and Canally, C. (2014). Tracking human activity and well-being in natural environments using wearable sensors and experien-ce sampling. Social Science & Medicine, 106, pp.83–92. doi:10.1016/j.socsci-med.2014.01.048.

Dzedzickis, A., Kaklauskas, A. and Bucinskas, V. (2020). Human Emotion Rec-ognition: Review of Sensors and Methods. Sensors (Basel, Switzerland), 20(3), p.592. doi:10.3390/s20030592.

Eakins, J.P. (1996). Automatic image content retrieval-are we getting anywhere?.In proceedings of the third international conference on electronic library and visual information research pp. 123-135. Retrieved From https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Automatic+image+content+retrieval-are+we+getting+anywhere&btnG=.

Egger, M., Ley, M. and Hanke, S. (2019). Emotion Recognition from Physiolo-gical Signal Analysis: A Review. Electronic Notes in Theoretical Computer Science, 343, pp.35–55. doi:10.1016/j.entcs.2019.04.009.

Ekman, P. (1992). An argument for basic emotions. Cognition and Emotion, 6(3-4), pp.169–200. doi:10.1080/02699939208411068.

Ekman, P. (1993). Facial expression and emotion. American psychologist, 48(4), p.384.

Ernawati, J., Adhitama, M.S., Surjono and Sudarmo, B.S. (2016). Urban DesignQualities Related Walkability in a Commercial Neighbourhood. Environment-Behaviour Proceedings Journal, 1(4), p.242. doi:10.21834/e-bpj.v1i4.385.

Ewing, R. and Handy, S. (2009). Measuring the Unmeasurable: Urban Design Qualities Related to Walkability. Journal of Urban Design, 14(1), pp.65–84. doi:10.1080/13574800802451155.

Fathullah, A. and Willis, K. (2018). Engaging the Senses: The Potential of Emotional Data for Participation in Urban Planning. Urban Science, 2(4), pp.98. doi:10.3390/urbansci2040098.

Ferreira BRA. (2018). Emotions recognition based on sensor fusion and machi-ne learning techniques [Master Thesis]. Universidade do Minho. Portugal; Available: http://www3.dsi.uminho.pt/pimenta/supmsdsis/bd/PT4ILMPFOKCQU/ProjDissertacaoBrunoFerreiraA72285.pdf

Fischer, K.W., Shaver, P.R. and Carnochan, P. (1990). How Emotions Develop and How they Organise Development. Cognition & Emotion, 4(2), pp.81–127.doi:10.1080/02699939008407142.

Gehl, J. and Gemzøe, L. (2001). New city spaces. The Danish Architectural Press.

Geiser, M. and Walla, P. (2011). Objective Measures of Emotion During Virtu-al Walks through Urban Environments. Applied Sciences, 1(1), pp.1–11. doi:10.3390/as1010001.

Girardi, D., Lanubile, F. and Novielli, N. (2017). Emotion detection using non-invasive low cost sensors. 2017 Seventh International Conference on Affecti-ve Computing and Intelligent Interaction (ACII). doi:10.1109/acii.2017.8273589.

Gjerde, M. Visual Aesthetic Perception and Judgement Of Urban Streets-capes. Available at: https://www.kenken.go.jp/japanese/contents/cib/w101/pdf/m-tg/1005salford/session03.pdf [Accessed 7 Oct. 2022].

Golder, S.A. and Macy, M.W. (2011). Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. Science, 333(6051), pp.1878–1881. doi:10.1126/science.1202775.

Gong, P., Ma, H.T. and Wang, Y. (2016). Emotion recognition based on the multiple physiological signals. 2016 IEEE International Conference on Real-ti-me Computing and Robotics (RCAR). pp. 140-143. doi:10.1109/rcar.2016.7784015.

Goshvarpour, A., Abbasi, A. and Goshvarpour, A. (2017). An accurate emotion recognition system using ECG and GSR signals and matching pursuit metho-

d. Biomedical Journal, 40(6), pp.355–368. doi:10.1016/j.bj.2017.11.001.

Gu, Y., Tan, S.L., Wong, K.J., Ho, M.H.R. and Qu, L. (2009). Using GA-bas-ed feature selection for emotion recognition from physiological signals. 2008

International Symposium on Intelligent Signal Processing and CommunicationsSystems. doi:10.1109/ispacs.2009.4806747.

Qingfeng Guan, (2019). GPU-CUDA-enabled Semantic Segmentation App (v1.0). [Software]. [Accessed 2020.12.20]. Retrieved From: https://www.urbancomp.n-et/archives/semantic-segmentation-software-for-visual-images-based-on-fcn.

Harvey, C. and Aultman-Hall, L. (2015). Measuring Urban Streetscapes for Li-vability: A Review of Approaches. The Professional Geographer, 68(1), pp.149–158. doi:10.1080/00330124.2015.1065546.

Hassani, S., Bafadel, I., Bekhatro, A., Al Blooshi, E., Ahmed, S. and Alahmad,M. (2017). Physiological signal-based emotion recognition system. 2017 4th IEEE International Conference on Engineering Technologies and Applied Sci-ences (ICETAS). doi:10.1109/icetas.2017.8277912.

Healey, J.A. and Picard, R.W. (2005). Detecting Stress During Real-World Dri-ving Tasks Using Physiological Sensors. IEEE Transactions on Intelligent Tr-ansportation Systems, 6(2), pp.156–166. doi:10.1109/tits.2005.848368.

Heckscher, A. and Robinson, P.C. (1977). Open spaces: The life of American cities. HarperCollins Publishers.

Hernandez, J., McDuff, D., Fletcher, R. and Picard, R.W., (2013). March. Inside-out: Reflecting on your inner state. In 2013 IEEE International Conference on Pervasive Computing and Communications Workshops. pp. 324-327.

Hillier, B. and Penn, A. (1994). Virtuous circles, building sciences and the sci-ence of buildings: using computers to integrate product and process in the built environment. Design Studies, 15(3), pp.332–365. doi:10.1016/0142-694x(94)90017-5.

Hockenbury, D.H. and Hockenbury, S.E. (ed.) (2014). Discoovering psychology. 6th edn. New York: Worth Publishers.

Hogertz, C. (2010) Emotions of urban pedestrians: Sensory mapping. Pedestrian-s' quality needs, Zurich, Switzerland: European Co-operation in Science and Technology. PQN Final Report 31, pp.31-52.

Huang, H., Klettner, S., Schmidt, M., Gartner, G., Leitinger, S., Wagner, A., a-nd Steinmann, R. (2014) AffectRoute: Considering people's affective responses to environments for enhancing route-planning services. International Journal of Geographical Information Science 28(12): 2456-2473.

Hui, T. K. and Sherratt, R. S. (2018). Coverage of emotion recognition for common wearable biosensors. Biosensors, 8(2), p.30. https://doi.org/10.3390/bios8020030.

Iacoviello, D., Petracca, A., Spezialetti, M. and Placidi, G. (2016). A Classific-ation Algorithm for Electroencephalography Signals by Self-Induced Emotion-al Stimuli. IEEE Transactions on Cybernetics, 46(12), pp.3171–3180. doi:10.1109/tcyb.2015.2498974.

Iliou, T. and Anagnostopoulos, C.-N. (2009). Comparison of Different Classifie-rs for Emotion Recognition. 2009 13th Panhellenic Conference on Informatic-s. doi:10.1109/pci.2009.7.

Jahanmohan, T. (2016). Identifying and measuring urban design qualities relatedto walkability – special reference to jaffna down town. dl.lib.uom.lk. Available at: http://dl.lib.uom.lk/handle/123/12958 [Accessed 12 Oct. 2022].

James, W. (1984). What is emotion? Mind, pp. 188-205.

Jang, E.-H., Park, B.-J., Park, M.-S., Kim, S.-H. and Sohn, J.-H. (2015). Anal-ysis of physiological signals for recognition of boredom, pain, and surprise emotions. Journal of Physiological Anthropology, 34(1). doi:10.1186/s40101-015-0063-5.

Jerritta, S., Murugappan, M., Nagarajan, R. and Wan, K. (2011). Physiological signals based human emotion Recognition: a review. IEEE Xplore. doi:10.1109/CSPA.2011.5759912.

Kalimeri, K. and Saitis, C. (2016). Exploring multimodal biosignal features for stress detection during indoor mobility. Proceedings of the 18th ACM Intern-ational Conference on Multimodal Interaction. pp.53-60. doi:10.1145/2993148.2993159.

Kalivoda, O., Vojar, J., Skřivanová, Z., and Zahradník, D. (2014). Consensus inlandscape preference judgements: The effects of landscape visual aesthetic quality and respondents' characteristics. Journal of Environmental Managemen-t 137: 36-44.

Kanjo, E., Al-Husain, L. and Chamberlain, A. (2015). Emotions in context: ex-amining pervasive affective sensing systems, applications, and analyses. Pers-onal and Ubiquitous Computing, 19(7), pp.1197–1212. doi:10.1007/s00779-015-0842-3.

Kanjo, E., Younis, E.M.G. and Sherkat, N. (2018). Towards unravelling the rel-ationship between on-body, environmental and emotion data using sensor inf-ormation fusion approach. Information Fusion, 40, pp.18–31. doi:10.1016/j.inffus.2017.05.005.

Kanjo, E., Younis, E.M.G. and Ang, C.S. (2019). Deep learning analysis of m-obile physiological, environmental and location sensor data for emotion dete-ction. Information Fusion, 49, pp.46–56. doi:10.1016/j.inffus.2018.09.001.

Keelawat, P., Thammasan, N., Numao, M. and Kijsirikul, B. (2021). A Compa-rative Study of Window Size and Channel Arrangement on EEG-Emotion R-ecognition Using Deep CNN. Sensors, 21(5), p.1678. doi:10.3390/s21051678.

Kemper, T.D. and Lazarus, R.S. (1992). Emotion and Adaptation. Contemporary Sociology, 21(4), p.522. doi:10.2307/2075902.

Khezri, M., Firoozabadi, M. and Sharafat, A.R. (2015). Reliable emotion recog-nition system based on dynamic adaptive fusion of forehead biopotentials and physiological signals. Computer Methods and Programs in Biomedicine, 122(2), pp.149–164. doi:10.1016/j.cmpb.2015.07.006.

Kim, K.H., Bang, S.W. and Kim, S.R. (2004). Emotion recognition system usi-ng short-term monitoring of physiological signals. Medical & Biological Eng-ineering & Computing, 42(3), pp.419–427. doi:10.1007/bf02344719.

Kołakowska, A., Szwoch, W. and Szwoch, M. (2020). A Review of Emotion Recognition Methods Based on Data Acquired via Smartphone Sensors. Sens-ors, 20(21), p.6367. doi:10.3390/s20216367.

Koohsari, M.J., Mavoa, S., Villanueva, K., Sugiyama, T., Badland, H., Kaczyn-ski, A.T., Owen, N. and Giles-Corti, B. (2015). Public open space, physical activity, urban design and public health: Concepts, methods and research ag-enda. Health & Place, 33, pp.75–82. doi:10.1016/j.healthplace.2015.02.009.

Kreibig, S.D. (2010). Autonomic nervous system activity in emotion: A revie-w. Biological Psychology, 84(3), pp.394–421. doi:10.1016/j.biopsycho.2010.03.010.

Kubios Oy. (2020). Kubios HRV Premium (ver. 3.4.3). [Software]. [Accessed 2020.04.15]. Retrieved from https://www.kubios.com/hrv-premium.

Lang, P.J., (1980). Self-assessment manikin. Gainesville, FL: The Center for Re-search in Psychophysiology, University of Florida.

Lazarus, R.S. (1982). Thoughts on the relations between emotion and cognitio-n. American Psychologist, 37(9), pp.1019–1024. doi:10.1037//0003-066x.37.9.1019.

Lazarus, R.S. (1991). Cognition and motivation in emotion. American Psycholo-gist, 46(4), pp.352–367. doi:10.1037/0003-066x.46.4.352.

Lazarus, R.S. (1995). Vexing Research Problems Inherent in Cognitive-Mediati-onal Theories of Emotion- and Some Solutions. Psychological Inquiry, 6(3), pp.183–196. doi:10.1207/s15327965pli0603_1.

Lee, B.J., Jang, T.Y., Wang, W. and Namgung, M. (2009). Design Criteria for an Urban Sidewalk Landscape Considering Emotional Perception. Journal of Urban Planning and Development, 135(4), pp.133–140. doi:10.1061/(asce)up.1943-5444.0000013.

Lee, S., Harada, A., & Stappers, P. J. (2002). Pleasure with products: Design based on Kansei. Pleasure with products: Beyond usability: 219-229.

Li, X., Hijazi, I., Koenig, R., Lv, Z., Zhong, C. and Schmitt, G. (2016). Asse-ssing Essential Qualities of Urban Space with Emotional and Visual Data B-ased on GIS Technique. ISPRS International Journal of Geo-Information, 5(11), p.218. doi:10.3390/ijgi5110218.

Li, W., Zhang, Z. and Song, A. (2021). Physiological-signal-based emotion rec-ognition: An odyssey from methodology to philosophy. Measurement, 172, p.108747. doi:10.1016/j.measurement.2020.108747.

Lin, F. C., She, S.J., Ngo, H.H., Dow, C.R. and Hsu, F.R., (2021). A Wearabl-e Embedded System for Assisting Cognition of Visually Impaired People by Street Scene Description. Journal of Computers, 32(1), pp.102-119.

Lisetti, C.L. and Nasoz, F. (2004). Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals. EURASIP Journal onAdvances in Signal Processing, 2004(11). pp.1-16. doi:10.1155/s1110865704406192.

Liu, L., Silva, E.A., Wu, C. and Wang, H. (2017). A machine learning-based method for the large-scale evaluation of the qualities of the urban environm-ent. Computers, Environment and Urban Systems, 65, pp.113–125. doi:10.1016/j.compenvurbsys.2017.06.003.

Londhe, S. and Borse, R. (2018). Emotion Recognition Based on Various Phys-iological Signals – A Review. ICTACT Journal on Communication Technolo-gy, 9(3), pp.1815–1822. doi:10.21917/ijct.2018.0265.

Marcus, G.E., Neuman, W.R. and MacKuen, M.B. (2015). Measuring EmotionalResponse: Comparing Alternative Approaches to Measurement. Political Scie-nce Research and Methods, 5(4), pp.733–754. doi:10.1017/psrm.2015.65.

Mauss, I.B. and Robinson, M.D. (2009). Measures of emotion: A review. Cogn-ition & Emotion, 23(2), pp.209–237. doi:10.1080/02699930802204677.

Mehmood, R.M. and Lee, H.J. (2016). A novel feature extraction method basedon late positive potential for emotion recognition in human brain signal patt-erns. Computers & Electrical Engineering, 53, pp.444–457. doi:10.1016/j.co-mpeleceng.2016.04.009.

Mehta, V. (2013). Evaluating Public Space. Journal of Urban Design, 19(1), pp.53–88. doi:10.1080/13574809.2013.854698.

Mehta, V. and Bosson, J.K. (2018). Revisiting Lively Streets: Social Interactio-ns in Public Space. Journal of Planning Education and Research, p.0739456X1878145. doi:10.1177/0739456x18781453.

Molavi, M., Yunus, J. bin and Akbari, E. (2012). Comparison of Different Me-thods for Emotion Classification. 2012 Sixth Asia Modelling Symposium. pp. 50–53. doi:10.1109/ams.2012.53.

Mossman, M.L. (2009). Religious and Philosophical Syncretization in Classical Chinese and Japanese Gardens. The International Journal of the Humanities: Annual Review, 7(5), pp.157–174. doi:10.18848/1447-9508/cgp/v07i05/42680.

Nasoz, F., Alvarez, K., Lisetti, C.L. and Finkelstein, N. (2004). Emotion recog-nition from physiological signals using wireless sensors for presence technologies. Cognition, Technology & Work, 6(1), pp.4–14. doi:10.1007/s10111-003-0143-x.

Nayak, S.K., Bit, A., Dey, A., Mohapatra, B. and Pal, K. (2018). A Review on the Nonlinear Dynamical System Analysis of Electrocardiogram Signal. Journal of Healthcare Engineering, 2018, pp.1–19. doi:10.1155/2018/6920420.

Nweke, H.F., Teh, Y.W., Mujtaba, G. and Al-garadi, M.A. (2019). Data fusion and multiple classifier systems for human activity detection and health moni-toring: Review and open research directions. Information Fusion, 46, pp.147–170. doi:10.1016/j.inffus.2018.06.002.

Ojha, V.K., Griego, D., Kuliga, S., Bielik, M., Buš, P., Schaeben, C., Treyer, L., Standfest, M., Schneider, S., König, R., Donath, D. and Schmitt, G. (2019). Machine learning approaches to understand the influence of urban enviro-nments on human's physiological response. Information Sciences, 474, pp.154–169. doi:10.1016/j.ins.2018.09.061.

Olsen AF, Torresen J. (2016). Smartphone accelerometer data used for detectinghumanemotions. In: The 3rd International Conference on Systems and Infor-matics (ICSAI 2016). Shanghai, China: IEEE; 2016. pp. 410–415. Available: http://doi.org/10.1109/ICSAI.2016.7810990

OriginLab Corporation. 2019. Origin (2019). [Software]. [Accessed 2020.04.22]. Retrieved from https://www.originlab.com/2019.

Oxford University Press. (2022). The Oxford Advanced American Dictionary. Available at: https://www.oxfordlearnersdictionaries.com/definition/american_english/emotion?q=emotion (Accessed: 8 October 2022)

Pallasmaa, J. (2014). Space, place and atmosphere. Emotion and peripherical perception in architectural experience. Lebenswelt. Aesthetics and philosophy of experience., (4). doi:10.13130/2240-9599/4202.

Peng, K.-C., Chen, T., Sadovnik, A. and Gallagher, A. (2015). A mixed bag of emotions: Model, predict, and transfer emotion distributions. IEEE Xplore. doi:10.1109/CVPR.2015.7298687.

Peng, K.-C., Sadovnik, A., Gallagher, A. and Chen, T. (2016). Where do emoti-ons come from? Predicting the Emotion Stimuli Map. IEEE Xplore. doi:10.1109/ICIP.2016.7532430.

Picard, R.W. (2000a). Toward computers that recognize and respond to user emotion. IBM Systems Journal, 39(3.4), pp.705–719. doi:10.1147/sj.393.0705.

Picard, R.W. (2000b). Affective computing. Cambridge, Massachusetts Etc.: Mit Press.

Picard, R.W. (2010). Emotion Research by the People, for the People. Emotion Review, 2(3), pp.250–254. doi:10.1177/1754073910364256.

Picard, R.W. and Scheirer, J. (2001). The galvactivator: A glove that senses and communicates skin conductivity. In Proceedings 9th Int. Conf. on HCI.

Picard, R.W., Vyzas, E. and Healey, J. (2001). Toward machine emotional intel-ligence: analysis of affective physiological state. IEEE Transactions on Patter-n Analysis and Machine Intelligence, 23(10), pp.1175–1191. doi:10.1109/34.954607.

Posner, J., Russell, J.A. and Peterson, B.S. (2005). The circumplex model of a-ffect: An integrative approach to affective neuroscience, cognitive developme-nt, and psychopathology. Development and Psychopathology, 17(03), pp.715–734. doi:10.1017/s0954579405050340.

Price, E.E., Caldwell, C.A. and Whiten, A. (2009). Comparative cultural cognit-ion. WIREs Cognitive Science, 1(1), pp.23–31. doi:10.1002/wcs.14.

Priya, D.T. and Udayan, J.D. (2020). Affective emotion classification using fea-ture vector of image based on visual concepts. The International Journal of Electrical Engineering & Education, p.002072092093683. doi:10.1177/0020720920936834.

Rahman, N.A., Ghani, I., Teh, M.Z. and Ibrahim, K.A., (2020). Rethinking urb-an public space: Physical and functional analysis through visual surveys. In IOP Conference Series: Earth and Environmental Science (Vol. 409, No. 1, p. 012002). IOP Publishing.

Resch, B., Puetz, I., Bluemke, M., Kyriakou, K. and Miksch, J. (2020). A Int-erdisciplinary Mixed-Methods Approach to Analyzing Urban Spaces: The Ca-se of Urban Walkability and Bikeability. International Journal of Environmen-tal Research and Public Health, 17(19), p.6994. doi:10.3390/ijerph17196994.

Resch, B., Summa, A., Sagl, G., Zeile, P. and Exner, J.-P. (2014). Urban Emo-tions—Geo-Semantic Emotion Extraction from Technical Sensors, Human Se-nsors and Crowdsourced Data. Lecture Notes in Geoinformation and Cartogr-aphy, pp.199–212. doi:10.1007/978-3-319-11879-6_14.

Roberts, H., Sadler, J. and Chapman, L. (2019). The value of Twitter data for determining the emotional responses of people to urban green spaces: A case study and critical evaluation. Urban studies 56(4): 818-835.

Rossetti, T., Lobel, H., Rocco, V., and Hurtubia, R. (2019) Explaining subjecti-ve perceptions of public spaces as a function of the built environment: A massive data approach. Landscape and Urban Planning 181, pp. 169-178.

Russell, J.A. and Mehrabian, A. (1978). Approach-avoidance and affiliation as functions of the emotion-eliciting quality of an environment. Environment andbehavior, 10(3), pp.355-387.

Russell, J.A. (1979). Affective space is bipolar. Journal of personality and soci-al psychology, 37(3), pp.345-356.

Russell, J.A. (2003). Core affect and the psychological construction of emotion-

n. Psychological Review, 110(1), pp.145–172. doi:10.1037/0033-295x.110.1.145.

Russell, J. A., & Mehrabian, A. (1978). Approach-Avoidance and Affiliation as Functions of the Emotion-Eliciting Quality of an Environment. Environment and Behavior, 10(3), 355–387.

Schachter, S. and Singer, J. (1962). Cognitive, social, and physiological determ-inants of emotional state. Psychological Review, 69(5), pp.379–399. doi:10.1037/h0046234.

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M.and Scherer, K.R. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. PLOS ONE, 12(6), p.e0178899. doi:10.1371/journal.pone.0178899.

Schneider, S., Fröhlich, J., Bielik, M. and König, R. (2014). Space Synthesizer-An experimental setup for investigating the relationships between urban formand emotional responses. Design Cognition and Behavior: Usability in the Built Environment, p.23.

Shoval, N., Schvimer, Y., and Tamir, M. (2018) Real-time measurement of tou-rists' objective and subjective emotions in time and space. Journal of Travel Research 57(1): 3-16.

Shoval, N., Schvimer, Y., and Tamir, M. (2018) Tracking technologies and urb-an analysis: Adding the emotional dimension. Cities 72: 34-42.

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X. and Yang, X. (2018). A Review of Emotion Recognition Using Physiological Signals. Sensors, 18(7), pp.103-126. doi:10.3390/s18072074.

Shukla, J., Barreda-Angeles, M., Oliver, J., Nandi, G.C. and Puig, D. (2019). Feature Extraction and Selection for Emotion Recognition from ElectrodermalActivity. IEEE Transactions on Affective Computing, pp. 857–869. doi:10.1109/taffc.2019.2901673.

Shumailov, I. and Gunes, H. (2017). Computational analysis of valence and arousal in virtual reality gaming using lower arm electromyograms. 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII). doi:10.1109/acii.2017.8273595.

Steinmetz-Wood, M., Velauthapillai, K., O'Brien, G., and Ross, N. A. (2019) Assessing the micro-scale environment using Google Street View: The virtualsystematic tool for evaluating pedestrian streetscapes. BMC Public Health 19(1): 1246. https://doi.org/10.1186/s12889-019-7460-3.

Tan, C., Ceballos, G., Kasabov, N. and Puthanmadam Subramaniyam, N. (2020). FusionSense: Emotion Classification Using Feature Fusion of Multimodal Data and Deep Learning in a Brain-Inspired Spiking Neural Network. Senso-rs, 20(18), p.5328. doi:10.3390/s20185328.

Tang, J. and Long, Y. (2019) Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing. Landscape and Urban Planning, 191: 103436. https://doi.org/10.1016/j.landurbplan.2018.09.015.

Taylor, B.T., Fernando, P., Bauman, A.E., Williamson, A., Craig, J.C. and Redman, S. (2011). Measuring the Quality of Public Open Space Using Google Earth. American Journal of Preventive Medicine, 40(2), pp.105–112. doi:10.1016/j.amepre.2010.10.024.

Torres-Valencia, C., Álvarez-López, M. and Orozco-Gutiérrez, Á. (2016). SVM-based feature selection methods for emotion recognition from multimodal data. Journal on Multimodal User Interfaces, 11(1), pp.9–23. doi:10.1007/s12193-016-0222-y.

Uebersax, J.S. (1982). A Generalized Kappa Coefficient. Educational and Psychological Measurement, 42(1), pp.181–183. doi:10.1177/0013164482421018.

Vallotton, C.D. (2008). Signs of emotion: What can preverbal children ‘say’ about internal states?. Infant Mental Health Journal, 29(3), pp.234–258. doi:10.1002/imhj.20175.

Vergouwe, Y., Steyerberg, E.W., Eijkemans, M.J.C. and Habbema, J.D.F. (2005).Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. Journal of Clinical Epidemiology, 58(5), pp.475–483. doi:10.1016/j.jclinepi.2004.06.017.

Verma, G.K. and Tiwary, U.S. (2014). Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiologi-cal signals. NeuroImage, 102, pp.162–172. doi:10.1016/j.neuroimage.2013.11.007.

Wagner, J., Jonghwa Kim and Andre, E. (2005). From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extra-ction and Classification. 2005 IEEE International Conference on Multimedia and Expo. doi:10.1109/icme.2005.1521579.

Wang, S. and Wang, X. (2005). Emotion Semantics Image Retrieval: An Brief Overview. Affective Computing and Intelligent Interaction, pp.490–497. doi:10.1007/11573548_63.

Weber, R., Schnier, J. and Jacobsen, T. (2008). Aesthetics of Streetscapes: Infl-uence of Fundamental Properties on Aesthetic Judgments of Urban Space. Perceptual and Motor Skills, 106(1), pp.128–146. doi:10.2466/pms.106.1.128-146.

Weiss, C.C., Purciel, M., Bader, M., Quinn, J.W., Lovasi, G., Neckerman, K.M.and Rundle, A.G. (2011). Reconsidering Access: Park Facilities and Neigh- borhood Disamenities in New York City. Journal of Urban Health, 88(2), pp.297–310. doi:10.1007/s11524-011-9551-z.

Wen, W., Liu, G., Cheng, N., Wei, J., Shangguan, P. and Huang, W. (2014). Emotion

Recognition Based on Multi-Variant Correlation of Physiological Si-gnals. IEEE Transactions on Affective Computing, 5(2), pp.126–140. doi:10.1109/taffc.2014.2327617.

Whyte, W. H. (1980) The Social Life of Small Urban Space [Motion Pictur-e]. Santa Monica, CA: Direct Cinema Limited.

Wilhelm, F. H. and Grossman, P. (2010) Emotions beyond the laboratory: The-oretical fundaments, study design, and analytic strategies for advanced ambu-latory assessment. Biological Psychology 84(3): 552-569.

William, H. W (1980). The social life of small urban spaces. New York: Proje-ct For Public Spaces.

Woolley, H. (2003). Urban open spaces. Taylor & Francis.

Xiang, L. and Papastefanou, G. (2019). Isovist and Psycho-Physiological Stress at the Pedestrian Level: A Real-Time Measurement Case Study in a High- Density City. In IS THIS THE REAL WORLD? Perfect Smart Cities vs. R-eal Emotional Cities. Proceedings of REAL CORP 2019, 24th International Conference on Urban Development, Regional Planning and Information Soci-ety (pp. 463-471).

Yamamoto, K., Yoshikawa, T., & Furuhashi, T. (2005). Division method of su-bjects by individuality for stratified analysis of SD evaluation data. In IEEE International Symposium on Micro-NanoMechatronics and Human Science: 29-34.

Yao, W., Zhang, X. and Gong, Q. (2021). The effect of exposure to the natur-al environment on stress reduction: A meta-analysis. Urban Forestry & UrbanGreening, 57, p.126932. doi:10.1016/j.ufug.2020.126932.

Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., Wang, R., Wang, J. and Guan, Q. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. International Journal ofGeographical Information Science, 33(12), pp.2363–2384. doi:10.1080/13658816.2019.1643024.

Yoshikawa, A. (2000). Subjective information processing: Its foundation and applications. International Journal of Biomedical Soft Computing and Human Sciences: the official journal of the Biomedical Fuzzy Systems Association, 6(1), pp.75-83.

You, Q., Luo, J., Jin, H. and Yang, J. (2016). Building a Large Scale Dataset for Image Emotion Recognition: The Fine Print and The Benchmark. Proceedings of the AAAI Conference on Artificial Intelligence, 30(1). doi:10.1609/aaai.v30i1.9987.

Zadra, J.R. and Clore, G.L. (2011). Emotion and perception: the role of affecti-ve information. Wiley Interdisciplinary Reviews: Cognitive Science, 2(6), pp.676–685. doi:10.1002/wcs.147.

Zeile, P., Resch, B., Dörrzapf, L., Exner, J. P., Sagl, G., Summa, A., and Sud-manns, M. (2015) Urban emotions– Tools of integrating people’s perception into urban planning.

In From Vision to Reality for Vibrant Cities and Regio-ns, Proceedings of the 20th International Conference on Urban Planning, Re-gional Development, and Information Society; 5-7 May, Ghent, Belgium: pp. 905-912.

Zepf, S., Hernandez, J., Schmitt, A., Minker, W. and Picard, R.W. (2021). Dri-ver Emotion Recognition for Intelligent Vehicles. ACM Computing Surveys, 53(3), pp.1–30. doi:10.1145/3388790.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H. and Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machi-ne learning. Landscape and Urban Planning, 180, pp.148–160. doi:10.1016/j. landurbplan.2018.08.020.

Zhang, H., Augilius, E., Honkela, T., Laaksonen, J., Gamper, H. and Alene, H. (2011). Analyzing Emotional Semantics of Abstract Art Using Low-Level Image Features. Advances in Intelligent Data Analysis X, pp.413–423. doi:10.1007/978-3-642-24800-9_38.

Zhang, J., Yin, Z., Chen, P. and Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. Information Fusion, 59, pp.103-126.

Zhang, X., Xu, C., Xue, W., Hu, J., He, Y. and Gao, M. (2018). Emotion Recognition Based on Multichannel Physiological Signals with Comprehensive Nonlinear Processing. Sensors, 18(11), p.3886. doi:10.3390/s18113886.

Zhao, C. S., Matsumoto, N., Liu, T. F., Yuan, Y. Q. and Kawata, K. (2003). A Study on the Image of Landscape of Japanese and Chinese Gardens. Journal of Asian Architecture and Building Engineering, 2(2), pp.b69–b76. doi:10.3130/jaabe.2.b69.

Zhao, S. (2016). Image Emotion Computing. Proceedings of the 24th ACM international conference on Multimedia. doi:10.1145/2964284.2971473.

Zhao, S., Yao, H., Gao, Y., Ding, G. and Chua, T.-S. (2018). Predicting Perso-nalized Image Emotion Perceptions in Social Networks. IEEE Transactions onAffective Computing, 9(4), pp.526–540. doi:10.1109/taffc.2016.2628787.

Zhou, Q. (2017). Multi-layer affective computing model based on emotional psychology. Electronic Commerce Research, 18(1), pp.109–124. doi:10.1007/s10660-017-9265-8.

Appendix

Appendix A: Valence statistics of participants in 10 public spaces.

JAIST campus

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	1	2	1	1	2	2	-1	1	1	2	1	1	1	2	0	1	1	1	-1	0	1	2	1	0	0	-2	2	1	2	2	0	1	-1	-1	0	2

The residential area in Yokaichi

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	0	-1	2	-1	0	0	1	0	1	1	1	0	0	1	2	2	1	2	1	0	0	0	0	0	1	0	1	0	1	0	1	1	0	0	1	1

Kenroku-en

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	0	1	1	2	1	0	0	2	1	0	1	0	0	1	2	1	1	1	1	1	1	1	0	0	0	0	1	0	0	1	1	0	1	1	1	1

D. T. Suzuki Museum

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	0	2	2	2	1	-1	2	1	0	2	0	0	-1	0	0	1	0	1	2	0	0	1	2	0	0	1	1	0	0	0	0	0	1	1	1	1

The Higashi Chaya District in Kanazawa

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	0	1	2	1	1	0	1	0	-1	2	0	0	1	0	2	2	0	-1	0	2	1	0	0	-1	1	1	1	1	0	0	0	0	1	1	2	1

DLPU Campus

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09				P10				P11			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	0	1	2	1	0	1	2	-1	0	1	1	1	0	0	2	-1	2	1	0	0	1	0	1	2	1	0	1	1	0	1	1	1	0	-1	1	-1	1	-1	0	-1	1	0	1	0

Meilin Park residential area

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09				P10				P11			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	0	1	0	-1	-1	-1	0	0	0	0	1	-1	0	0	0	0	-1	1	-1	-1	-1	0	0	1	0	1	1	-1	-1	1	-1	0	0	0	1	-1	-1	0	-1	1	0	0	0	0

Huarun 24 City Park

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09				P10				P11			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	1	2	1	1	-1	0	2	0	0	1	1	0	0	1	2	1	0	0	1	0	1	1	2	0	0	1	2	1	0	1	2	0	0	1	2	1	0	1	2	0	1	1	1	1

Dalian Heroes Memorial Park

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09				P10				P11			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	-1	0	-2	1	0	1	0	0	1	1	1	1	2	0	1	1	-1	-1	1	0	0	1	1	0	1	1	0	-1	0	1	0	1	0	-1	0	1	-2	-1	0	2	1	1	1	1

Dalian Ganjingli · Dongshi historic district

Participants	P01				P02				P03				P04				P05				P06				P07				P08				P09				P10				P11			
Sections	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Valences	0	1	1	1	1	1	1	0	0	0	1	0	0	-1	1	-1	1	0	1	0	0	1	1	-1	0	1	2	-1	-1	-1	0	-1	1	-1	1	0	0	-1	2	0	0	1	1	0

Appendix B: Output of external validation and classification accuracy.

Binary	LR		DT C5.0		NN	
	Predicted	**Actual**	**Predicted**	**Actual**	**Predicted**	**Actual**
	1	1	1	1	1	1
	-1	1	1	1	1	1
	1	1	1	1	1	1
	-1	1	1	1	1	1
	-1	1	-1	1	-1	1
	-1	1	1	1	-1	1
	1	1	-1	1	1	1
	-1	1	-1	1	1	1
	-1	1	1	1	-1	1
	1	1	1	1	1	1
	-1	1	1	1	-1	1
	1	1	1	1	1	1
	-1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	-1	1
	1	1	1	1	1	1
	-1	1	1	1	-1	1
	1	1	1	1	-1	1
	-1	1	1	1	1	1
	-1	1	1	1	-1	1
	1	1	1	1	-1	1
	-1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	-1	1	1	1
	-1	1	-1	1	-1	1
	1	1	1	1	1	1
	1	-1	1	-1	1	-1
	1	1	1	1	1	1
	1	-1	1	-1	-1	-1
	-1	1	1	1	1	1
	1	1	1	1	1	1
	-1	1	-1	1	-1	1
	-1	1	-1	1	-1	1
	-1	-1	-1	-1	-1	-1
	-1	1	-1	1	-1	1
	1	1	1	1	1	1
	-1	-1	1	-1	-1	-1
	1	-1	-1	-1	1	-1
	-1	-1	-1	-1	1	-1
	1	-1	1	-1	-1	-1
	1	1	-1	1	-1	1
	-1	-1	1	-1	-1	-1
	-1	1	1	1	-1	1
	1	-1	1	-1	-1	-1
	-1	1	-1	1	1	1
	1	1	1	1	-1	1
	1	1	1	1	1	1
	46.80%	**47**	**65.90%**	**47**	**61.70%**	**47**

Ternary	LR		DT C5.0		NN	
	Predicted	**Actual**	**Predicted**	**Actual**	**Predicted**	**Actual**
	1	0	0	0	1	0
	-1	1	-1	1	-1	1
	1	1	1	1	1	1
	0	1	1	1	1	1
	0	1	0	1	0	1
	0	0	1	0	1	0
	-1	0	-1	0	0	0
	1	1	1	1	1	1
	-1	1	-1	1	1	1
	0	0	0	0	0	0
	1	1	1	1	1	1
	0	0	0	0	1	0
	-1	1	-1	1	-1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	0	1	0	1	0	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	0	1	0	1	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	1	1	1	1	1	1
	0	1	1	1	1	1
	1	0	1	0	0	0
	1	1	1	1	1	1
	1	1	1	1	0	1
	0	1	0	1	0	1
	1	1	1	1	1	1
	-1	0	-1	0	-1	0
	1	1	1	1	1	1
	-1	1	1	1	1	1
	0	1	0	1	0	1
	1	1	1	1	1	1
	0	1	0	1	0	1
	1	1	1	1	1	1
	-1	0	-1	0	-1	0
	0	0	0	0	0	0
	-1	0	-1	0	0	0
	1	1	1	1	1	1
	1	0	1	0	1	0
	0	0	0	0	0	0
	1	-1	1	-1	1	-1
	0	1	0	1	0	1
	1	-1	1	-1	1	-1
	1	1	1	1	1	1
	0	0	0	0	0	0
	1	1	1	1	1	1
	1	0	0	0	0	0
	1	0	0	0	0	0
	1	1	1	1	1	1
	0	1	0	1	0	1
	0	-1	0	-1	0	-1
	-1	0	-1	0	-1	0
	1	1	1	1	1	1
	-1	1	-1	1	1	1
	1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1
	0	-1	0	-1	0	-1
	-1	0	-1	0	-1	0
	1	-1	1	-1	1	-1
	-1	1	-1	1	-1	1
	-1	-1	-1	-1	-1	-1
	0	1	0	1	0	1
	0	0	0	0	0	0
	-1	0	0	0	0	0
	-1	-1	-1	-1	-1	-1
	-1	1	-1	1	1	1
	-1	0	-1	0	-1	0
	1	0	1	0	1	0
	0	1	0	1	0	1
	-1	1	-1	1	-1	1
	0	0	1	0	1	0
	44.40%	**72**	**53.30%**	**72**	**56.90%**	**72**

Quinary	**LR**		**DT C5.0**		**NN**	
	Predicted	**Actual**	**Predicted**	**Actual**	**Predicted**	**Actual**
	2	0	2	0	-1	0
	1	1	-1	1	2	1
	2	1	1	1	2	1
	2	2	1	2	0	2
	-1	1	1	1	1	1
	-1	0	2	0	0	0
	2	0	2	0	2	0
	1	2	2	2	2	2
	2	1	1	1	2	1
	0	0	0	0	1	0
	2	1	2	1	1	1
	0	0	0	0	0	0
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	0	1	0	1
	-1	1	1	1	1	1
	-1	1	2	1	0	1
	1	0	1	0	1	0
	0	0	0	0	0	0
	1	0	0	0	1	0
	-1	1	1	1	-1	1
	2	1	1	1	1	1
	1	0	1	0	1	0
	1	1	2	1	2	1
	2	1	1	1	2	1
	1	1	2	1	1	1
	1	1	1	1	-1	1
	1	0	1	0	1	0
	-1	1	-1	1	1	1
	1	1	1	1	1	1
	0	1	-1	1	1	1
	1	1	1	1	0	1
	1	1	2	1	1	1
	1	1	0	1	0	1
	-1	0	-1	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	1	1	2	1	1	1
	-1	0	-1	0	0	0
	1	0	0	0	1	0
	-1	-1	1	-1	1	-1
	0	1	1	1	1	1
	-1	-1	-1	-1	1	-1
	-1	1	0	1	1	1
	1	0	0	0	-1	0
	-1	1	1	1	0	1
	0	0	0	0	0	0
	1	0	2	0	0	0
	1	1	1	1	0	1
	0	1	0	1	0	1
	0	-1	0	-1	-1	-1
	1	0	0	0	0	0
	1	1	1	1	-1	1
	0	2	-1	2	2	2
	-1	-1	0	-1	0	-1
	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1
	1	0	-1	0	0	0
	0	-1	-1	-1	-1	-1
	-1	1	-1	1	1	1
	-1	-1	-1	-1	0	-1
	-1	1	0	1	2	1
	-1	0	-1	0	2	0
	0	0	1	0	0	0
	2	-1	-1	-1	-1	-1
	2	2	-1	2	2	2
	-1	0	0	0	-1	0
	2	0	0	0	0	0
	1	1	-1	1	1	1
	2	1	1	1	1	1
	1	0	2	0	0	0
	43.10%	**72**	**51.40%**	**72**	**55.60%**	**72**

Binary	DT C5.0（Boosting）		RF（Bagging）		NN（Boosting）	
	Predicted	**Actual**	**Predicted**	**Actual**	**Predicted**	**Actual**
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	-1	1
	1	1	1	1	1	1
	-1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	-1	1	1	1	1	1
	-1	1	1	1	-1	1
	1	1	1	1	1	1
	-1	1	1	1	1	1
	-1	1	1	1	-1	1
	1	1	1	1	1	1
	-1	1	1	1	-1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	-1	1	1	1	1	1
	1	1	1	1	1	1
	1	-1	1	-1	-1	-1
	1	1	1	1	1	1
	1	-1	1	-1	-1	-1
	-1	1	1	1	1	1
	1	1	1	1	1	1
	-1	1	-1	1	-1	1
	1	1	1	1	-1	1
	-1	-1	-1	-1	-1	-1
	-1	1	-1	1	1	1
	1	1	1	1	1	1
	-1	-1	1	-1	-1	-1
	1	-1	-1	-1	1	-1
	-1	-1	1	-1	-1	-1
	1	-1	1	-1	-1	-1
	1	1	1	1	-1	1
	-1	-1	1	-1	-1	-1
	1	1	1	1	1	1
	1	-1	1	-1	-1	-1
	-1	1	-1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	65,90%	**47**	**78.70%**	**47**	**80.90%**	**47**

Ternary	DT C5.0（Boosting）		RF（Bagging）		NN（Boosting）	
	Predicted	**Actual**	**Predicted**	**Actual**	**Predicted**	**Actual**
	0	0	1	0	0	0
	-1	1	1	1	0	1
	1	1	1	1	1	1
	0	1	1	1	0	1
	0	1	1	1	1	1
	0	0	1	0	0	0
	-1	0	1	0	0	0
	1	1	1	1	0	1
	1	1	1	1	-1	1
	0	0	0	0	0	0
	1	1	1	1	1	1
	0	0	0	0	0	0
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	1	1	1	1
	1	1	0	1	0	1
	1	1	1	1	1	1
	1	1	0	1	0	1
	1	0	0	0	0	0
	0	0	0	0	0	0
	0	0	1	0	0	0
	1	1	1	1	1	1
	0	1	1	1	0	1
	1	0	1	0	1	0
	1	1	1	1	1	1
	1	1	1	1	1	1
	0	1	1	1	1	1
	1	1	1	1	1	1
	-1	0	0	0	1	0
	1	1	1	1	-1	1
	1	1	1	1	1	1
	0	1	0	1	-1	1
	1	1	1	1	1	1
	1	1	1	1	0	1
	1	1	1	1	1	1
	-1	0	-1	0	0	0
	0	0	0	0	1	0
	-1	0	-1	0	1	0
	1	1	0	1	1	1
	1	0	0	0	0	0
	0	0	1	0	0	0
	1	-1	1	-1	1	-1
	0	1	1	1	1	1
	1	-1	1	-1	1	-1
	0	1	1	1	1	1
	0	0	0	0	0	0
	1	1	1	1	0	1
	0	0	0	0	0	0
	1	0	0	0	1	0
	1	1	1	1	1	1
	0	1	0	1	1	1
	0	-1	-1	-1	-1	-1
	-1	0	0	0	0	0
	1	1	-1	1	0	1
	-1	1	0	1	1	1
	1	-1	0	-1	0	-1
	-1	-1	1	-1	1	-1
	0	-1	-1	-1	-1	-1
	-1	0	0	0	0	0
	-1	-1	-1	-1	-1	-1
	1	1	0	1	0	1
	-1	-1	-1	-1	-1	-1
	0	1	0	1	0	1
	0	0	0	0	1	0
	-1	0	1	0	-1	0
	-1	-1	1	-1	-1	-1
	-1	1	1	1	1	1
	-1	0	-1	0	-1	0
	0	0	0	0	0	0
	1	1	1	1	0	1
	1	1	0	1	1	1
	0	0	0	0	0	0
	54.20%	**72**	**65.30%**	**72**	**62.50%**	**72**

Quinary	DT C5.0 (Boosting)		RF (Baggingh)		NN(Boosting)	
	Predicted	**Actual**	**Predicted**	**Actual**	**Predicted**	**Actual**
	2	0	2	0	0	0
	1	1	1	1	2	1
	2	1	1	1	2	1
	2	2	2	2	2	2
	-1	1	1	1	2	1
	-1	0	2	0	0	0
	2	0	2	0	2	0
	2	2	2	2	2	2
	2	1	1	1	2	1
	0	0	0	0	0	0
	2	1	2	1	1	1
	0	0	0	0	0	0
	1	1	1	1	1	1
	1	1	2	1	1	1
	1	1	1	1	1	1
	1	1	0	1	0	1
	-1	1	1	1	1	1
	-1	1	2	1	0	1
	1	0	1	0	1	0
	0	0	0	0	0	0
	1	0	0	0	1	0
	1	1	1	1	1	1
	2	1	1	1	1	1
	1	0	1	0	1	0
	1	1	2	1	2	1
	2	1	1	1	2	1
	1	1	2	1	1	1
	1	1	1	1	1	1
	1	0	1	0	1	0
	-1	1	-1	1	1	1
	1	1	1	1	1	1
	0	1	-1	1	1	1
	1	1	1	1	0	1
	1	1	2	1	1	1
	1	1	0	1	0	1
	-1	0	-1	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	1	1	2	1	1	1
	-1	0	-1	0	0	0
	1	0	0	0	1	0
	-1	-1	-1	-1	1	-1
	1	1	1	1	1	1
	-1	-1	-1	-1	1	-1
	-1	1	0	1	1	1
	1	0	0	0	0	0
	-1	1	1	1	0	1
	0	0	0	0	0	0
	1	0	2	0	0	0
	1	1	1	1	0	1
	0	1	0	1	0	1
	0	-1	0	-1	-1	-1
	0	0	0	0	0	0
	1	1	1	1	-1	1
	0	2	-1	2	2	2
	-1	-1	0	-1	0	-1
	-1	-1	-1	-1	1	-1
	-1	-1	-1	-1	-1	-1
	1	0	-1	0	0	0
	0	-1	-1	-1	-1	-1
	-1	1	-1	1	1	1
	-1	-1	-1	-1	0	-1
	-1	1	0	1	1	1
	-1	0	-1	0	2	0
	0	0	0	0	0	0
	2	-1	-1	-1	-1	-1
	2	2	-1	2	2	2
	-1	0	0	0	-1	0
	2	0	0	0	0	0
	1	1	-1	1	1	1
	2	1	1	1	0	1
	1	0	2	0	0	0
	47.20%	72	**55.60%**	72	**61.10%**	72

Appendix C: Statistics of the physical features of each section in the five experiment sites in Japan.

Items	Features	JAIST campus				Residential area in Yokaichi				Kenroku-en				D. T. Suzuki Museum				Higashi Chaya District			
		Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4
Scale characteristics	Width between buildings (W_1)(m)	54	42	36	27	8	23	7	13	N/A	N/A	N/A	N/A	4	20	5	N/A	15	3	5	4
	Width between boundary trees (W_2)(m)	15	6	N/A	8	N/A	10	N/A	N/A	4	8	13	8	N/A	N/A	2	50	10	N/A	N/A	N/A
	Width between pavement's boundaries (W_3)(m)	9	2	4	4	4	7	3	5	2	5	5	5	4	1	1	2	7	3	5	3
	Height of building's boundary (H_1)(m)	13	14	16.5	15	7	6	7	9	N/A	N/A	N/A	N/A	3	4	2	N/A	3	7	7	7
	Height of boundary tree (H_2)(m)	6	7	1	2	N/A	3	N/A	N/A	9	10	11	6	N/A	N/A	6	10	10	N/A	N/A	N/A
	W_1/H_1	4	3	2	2	1	4	1	1	N/A	N/A	N/A	N/A	1	5	3	N/A	5	0	1	1
	W_2/H_2	2	1	N/A	3	N/A	3	N/A	N/A	0	1	1	1	N/A	N/A	0	5	1	N/A	N/A	N/A
Boundary characteristics	Plane (P), line (L)	L/P	L	N/A	N/A	P/L	P/L	P	P	L	L	L/P	L/P	P	P	P/L	L	L	P	P	P
	Vertical (V) or horizontal (H)	V	V	H/V	V/H	V	H/V	V	V	V	V	V/H	V/H	V	V	V	H/V	V	V	V	V
	Hard (H) and soft (S)	H/S	S	H/S	H/S	H	H/S	H/S	H	S	S	S	S	H	H/S	S	S	S	H	H	H
	Number of boundary layers	3/2	2	2/1	2/1	1/2	1/2	3	1/2	1	2	2	2	1	2	2/3	2	2	1	1	1
	Width of soft boundary (m)	20.6	13	0	4,0	0	1	3	1	20	10	30	10	N/A	25	11.6	20	6	N/A	N/A	N/A
	Permeability of both sides of the line of sight (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	3	2	1	1	4	4	4	3	2	1	1	2	5	2	2	2	1	5	5	5
	Relationship between spatial boundary and visual boundary (1-coincidence, 2-spatial boundary is larger than visual boundary, 3-visual boundary is larger thanspatial boundary)	3	3	3	3	1	3	1	3	3	3	3	2	1	3	3	3	3	1	1	1
Spatial continuity	Continuity of spatial boundary (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	2	1	5	4	2	5	2	4	2	1	2	2	2	2	2	1	2	2	1	3
	number of space types	3	1	4	3	2	4	1	3	2	2	3	2	2	1	3	1	2	2	2	2
	Percentage of main space length	69%	94%	57%	49%	67%	55%	84%	56%	55%	78%	76%	55%	49%	100%	63%	100%	80%	56%	87%	74%
Visual characteristics	Enclosure degree (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	3	1	5	2	2	3	2	3	2	3	4	2	1	2	1	4	4	1	2	3
	Introversion or extroversion (1-very introverted，2-moderately introverted， 3-average，4-moderately extroverted，5-very extroverted)	3	1	5	2	2	4	2	4	3	3	4	3	1	2	1	4	4	1	2	3
	Openness and hierarchy of space (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	2	4	1	3	3	2	3	2	4	3	1	2	4	4	4	2	2	4	3	3
Thematic characteristics	Main scene	Y	N	N	N	N	N	N	N	N	Y	Y	N	N	Y	N	N	N	N	N	N
	Theme of space	Y	N	N	N	N	N	N	N	N	N	Y	N	Y	Y	Y	N	N	N	Y	N
Element characteristics	Number of material types	5	3	7	5	4	6	4	7	3	4	4	4	3	4	3	4	5	3	3	4
	Number of elements	6	5	9	7	7	9	8	10	5	7	7	6	6	6	5	6	8	5	5	7

Appendix D: Statistics of the physical features of each section in the five experiment sites in Dalian.

Items	Features	DLPU Campus				Meilinyuan residential area				Huarun twenty-four city park				Dalian Heroes Memorial Park				Dalian ganjingli· dongshi District			
		Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4	Section 1	Section 2	Section 3	Section 4
Scale Features	Width between buildings (W_1)(m)	51.0	23.0	51.5	110.0	N/A	24.0	24.0	37.0	N/A	N/A	N/A	68.0	N/A	N/A	N/A	N/A	12.0	12.0	52.0	15.0
	Width between boundary trees (W_2)(m)	13.5	8.0	14.5	8.0	25.0	4.0	16.0	N/A	N/A	4.0	35.0	N/A	6.0	2.5	3.0	3.0	N/A	N/A	22.0	N/A
	Width between pavement's boundaries (W_3)(m)	9.0	6.0	6.0	6.0	20.0	3.0	16.0	37.0	26.0	4.0	32.0	8.0	5.0	2.0	2.0	2.0	10.0	8.0	5.0	12.0
	Height of building's boundary (H_1)(m)	14.0	18.0	17.0	4.0	N/A	15.0	15.0	15.0	N/A	N/A	N/A	12.0	N/A	N/A	N/A	N/A	10.0	7.0	12.0	11.0
	Height of boundary tree (H_2)(m)	7.0	4.5	8.0	6.0	4.0	6.0	4.0	4.0	N/A	6.0	5.0	3.0	13.0	4.5	3.5	6.0	N/A	N/A	2.5	N/A
	W_1/H_1	15.5	1.3	3.0	27.5	N/A	1.6	1.6	2.5	N/A	N/A	N/A	5.7	N/A	N/A	N/A	N/A	1.2	1.7	4.3	1.4
	W_2/H_2	1.9	1.8	1.8	1.3	6.2	0.7	4.0	N/A	N/A	0.7	7.0	N/A	0.5	0.6	0.9	0.5	N/A	N/A	8.8	N/A
Boundary characteristics	Plane (P), line (L)	L	P/L	P/L	L	P/L	P/L	L	P	P	P/L	L/P	L/P	L	L	P/L	L	L	P	P/L	L
	Vertical (V) or horizontal (H)	V	H/V	H/V	V	H/V	H/V	V	H	H/V	H/V	V/H	V/H	V	V	V/H	V	V	V	V/H	V
	Hard (H) and soft (S)	H/S	H/S	S	S	S	H/S	H/S	H	H/S	S	S	H/S	S	S	H/S	H/S	H	H	H/S	H
	Number of boundary layers	1	1	2/3	2/1	2	1/2	1	1	1	3	2	2	2	1	3	3	1	1	2	2
	Width of soft boundary (m)	4.0	3.0	5.5	8.0	3.0	3.0	2.0	0.0	0.0	8.0	28.0	4.0	2.5	1.5	3.0	5.0	0.0	0.0	2.0	N/A
	Permeability of both sides of the line of sight (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	2	3	2	3	3	4	2	1	1	3	2	3	5	5	3	3	5	4	2	5
	Relationship between spatial boundary and visual boundary (1-coincidence, 2-spatial boundary is larger than visual boundary, 3-visual boundary is larger thanspatial boundary)	3	3	3	3	3	1	1	1	3	1	3	3	1	1	3	3	1	1	3	1
Spatial continuity	Continuity of spatial boundary (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	2	3	2	1	3	2	1	2	2	1	2	1	1	1	1	1	1	1	4	1
	number of space types	2	3	2	1	2	2	1	1	3	3	2	2	2	1	1	1	1	1	3	1
	Percentage of main space length	46%	70%	62%	100%	83%	74%	100%	96%	42%	83%	76%	86%	95%	100%	100%	100%	95%	100%	42%	100%
	Enclosure degree (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	2	2	2	3	5	2	4	5	3	4	5	3	2	1	2	2	3	3	3	3
	Introversion or extroversion (1-very introverted, 2-moderately introverted, 3-average, 4-moderately extroverted, 5-very extroverted)	3	2	2	2	5	2	4	4	4	4	5	3	2	1	2	2	3	3	3	3
	Openness and hierarchy of space (1-very good, 2-good, 3-average, 4-bad, 5-very bad)	2	3	3	3	1	4	2	2	2	1	1	3	4	5	4	4	3	3	2	3
Thematic characteristics	Main scene	N	N	N	N	N	N	N	N	Y	Y	Y	N	Y	N	Y	Y	N	N	Y	N
	Theme of space	N	Y	Y	N	N	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Element characteristics	Number of material types	3	3	3	2	3	5	4	6	4	4	4	4	3	2	4	4	4	5	5	5
	Number of elements	3	5	5	3	5	5	5	8	6	5	6	5	6	2	5	6	4	6	7	4

Appendix E: Classification of objects for image semantic segmentation (150 classes).

0 | Unkown objects | 未知对象
1 | wall | 墙
2 | building; edifice | 建筑
3 | sky | 天空
4 | floor; flooring | 地板
5 | tree | 树
6 | ceiling | 天花板
7 | road; route | 路
8 | bed | 床
9 | windowpane; window | 窗
10 | grass | 草
11 | cabinet | 柜子
12 | sidewalk; pavement | 人行道
13 | person; individual; someone; somebody; mortal; soul | 人
14 | earth; ground | 地面
15 | door; double door | 门
16 | table | 桌子
17 | mountain; mount | 山
18 | plant; flora; plant life | 植物
19 | curtain; drape; drapery; mantle; pall | 窗帘
20 | chair | 椅子
21 | car; auto; automobile; machine; motorcar | 汽车
22 | water | 水
23 | painting; picture | 绘画
24 | sofa; couch; lounge | 沙发
25 | shelf | 架子
26 | house | 房子
27 | sea | 海
28 | mirror | 镜子
29 | rug; carpet; carpeting | 地毯
30 | field | 场地
31 | armchair | 扶手椅
32 | seat | 座位
33 | fence; fencing | 围栏
34 | desk | 桌子
35 | rock; stone | 岩石
36 | wardrobe; closet; press | 衣柜
37 | lamp | 灯
38 | bathtub; bathing tub; bath; tub | 浴缸

39 | railing; rail | 铁路
40 | cushion | 坐垫
41 | base; pedestal; stand | 底座
42 | box | 盒子
43 | column; pillar | 柱
44 | signboard; sign | 招牌
45 | chest of drawers; chest; bureau; dresser | 衣橱
46 | counter | 柜台
47 | sand | 沙子
48 | sink | 水槽
49 | skyscraper | 摩天大楼
50 | fireplace; hearth; open fireplace | 壁炉
51 | refrigerator; icebox | 冰箱
52 | grandstand; covered stand | 看台
53 | path | 路径
54 | stairs; steps | 楼梯
55 | runway | 跑道
56 | case; display case; showcase; vitrine | 展示柜
57 | pool table; billiard table; snooker table | 台球桌
58 | pillow | 枕头
59 | screen door; screen | 纱门
60 | stairway; staircase | 楼梯
61 | river | 河
62 | bridge; span | 桥
63 | bookcase | 书柜
64 | blind; screen | 百叶窗
65 | coffee table; cocktail table | 咖啡桌
66 | toilet; can; commode; crapper; pot; potty; stool; throne | 厕所等
67 | flower | 花
68 | book | 书
69 | hill | 山
70 | bench | 长椅
71 | countertop | 工作台面
72 | stove; kitchen stove; range; kitchen range; cooking stove | 炉等
73 | palm; palm tree | 棕榈
74 | kitchen island | 厨房空间
75 | computer; computing machine; computing device; data processor; electronic computer; information processing system | 电脑
76 | swivel chair | 旋转椅
77 | boat | 船
78 | bar | 酒吧
79 | arcade machine | 街机
80 | hovel; hut; hutch; shack; shanty | 小屋
81 | bus; autobus; coach; charabanc; double-decker; jitney; motorbus; motorcoach; omnibus;

passenger vehicle | 公交车等
82 | towel | 毛巾
83 | light; light source | 光; 光源
84 | truck; motortruck | 卡车
85 | tower | 塔
86 | chandelier; pendant; pendent | 吊灯
87 | awning; sunshade; sunblind | 遮篷
88 | streetlight; street lamp | 路灯
89 | booth; cubicle; stall; kiosk | 摊位等
90 | television receiver; television; television set; tv; tv set; idiot box; boob tube; telly; goggle box | 电视等
91 | airplane; aeroplane; plane | 飞机
92 | dirt track | 土路
93 | apparel; wearing apparel; dress; clothes | 服饰
94 | pole | 杆子
95 | land; ground; soil | 土地
96 | bannister; banister; balustrade; balusters; handrail | 栏杆
97 | escalator; moving staircase; moving stairway | 自动扶梯
98 | ottoman; pouf; pouffe; puff; hassock | 脚凳
99 | bottle | 瓶子
100 | buffet; counter; sideboard | 餐柜
101 | poster; posting; placard; notice; bill; card | 海报
102 | stage | 舞台
103 | van | 货车
104 | ship | 船
105 | fountain | 喷泉
106 | conveyer belt; conveyor belt; conveyer; conveyor; transporter | 输送带
107 | canopy | 遮篷
108 | washer; automatic washer; washing machine | 洗衣机
109 | plaything; toy | 玩具
110 | swimming pool; swimming bath; natatorium | 游泳池
111 | stool | 凳子
112 | barrel; cask | 桶
113 | basket; handbasket | 篮子
114 | waterfall; falls | 瀑布
115 | tent; collapsible shelter | 帐篷
116 | bag | 袋
117 | minibike; motorbike | 小型机车
118 | cradle | 摇篮
119 | oven | 烤箱
120 | ball | 球
121 | food; solid food | 食物
122 | step; stair | 台阶
123 | tank; storage tank | 槽;储罐

124 | trade name; brand name; brand; marque | 商标
125 | microwave; microwave oven | 微波炉
126 | pot; flowerpot | 花盆
127 | animal; animate being; beast; brute; creature; fauna | 动物
128 | bicycle; bike; wheel; cycle | 自行车
129 | lake | 湖
130 | dishwasher; dish washer; dishwashing machine | 洗碗机
131 | screen; silver screen; projection screen | 投影屏幕
132 | blanket; cover | 毯子
133 | sculpture | 雕塑
134 | hood; exhaust hood | 罩
135 | sconce | 壁式烛台
136 | vase | 花瓶
137 | traffic light; traffic signal; stoplight | 交通信号灯
138 | tray | 托盘
139 | ashcan; trash can; garbage can; wastebin; ash bin; ash-bin; ashbin; dustbin; trash barrel; trash bin | 垃圾桶
140 | fan | 风扇
141 | pier; wharf; wharfage; dock | 码头
142 | crt screen | 屏幕
143 | plate | 盘子
144 | monitor; monitoring device | 监控
145 | bulletin board; notice board | 布告牌
146 | shower | 淋浴
147 | radiator | 散热器
148 | glass; drinking glass | 玻璃;玻璃杯
149 | clock | 时钟
150 | flag | 旗

Appendix F: The output of the two-step clustering

(1) Distribution of features in five clusters—continuous variables

Clusters

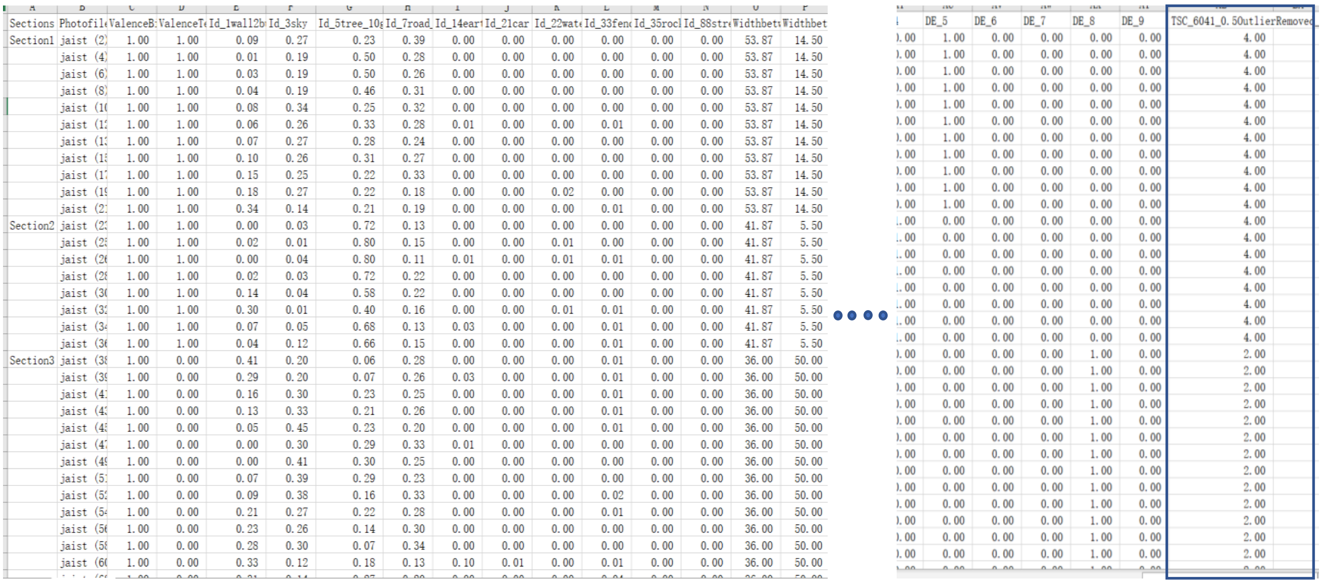
Input (Predictor) Importance

1.0 0.8 0.6 0.4 0.2 0.0

Cluster	4	1	5	3	2
Label					
Description					
Size	25.2% (100)	23.9% (95)	19.4% (77)	18.6% (74)	12.8% (51)
Inputs	Percentageofmainspacelength	Percentageofmainspacelength	Percentageofmainspacelength	Percentageofmainspacelength	Percentageofmainspacelength
	DS_1 0 (92.0%)	DS_1 0 (89.5%)	DS_1 1 (100.0%)	DS_1 0 (100.0%)	DS_1 0 (100.0%)
	DM_2 0 (70.0%)	DM_2 0 (100.0%)	DM_2 0 (100.0%)	DM_2 1 (100.0%)	DM_2 0 (100.0%)
	DL_2 0 (80.0%)	DL_2 0 (100.0%)	DL_2 0 (100.0%)	DL_2 0 (100.0%)	DL_2 1 (100.0%)
	DM_6 0 (100.0%)	DM_6 0 (100.0%)	DM_6 0 (100.0%)	DM_6 0 (100.0%)	DM_6 1 (72.5%)
	W2H2 2.30	W2H2 2.09	W2H2 0.75	W2H2 0.67	W2H2 19.89
	DL_3 1 (52.0%)	DL_3 1 (100.0%)	DL_3 0 (92.2%)	DL_3 0 (100.0%)	DL_3 0 (100.0%)
	DE_8 0 (100.0%)	DE_8 0 (100.0%)	DE_8 0 (100.0%)	DE_8 0 (100.0%)	DE_8 1 (66.7%)
	DS_4 0 (100.0%)	DS_4 0 (100.0%)	DS_4 0 (100.0%)	DS_4 0 (100.0%)	DS_4 1 (66.7%)
	Widthofsoftboundary 5.92	Widthofsoftboundary 17.72	Widthofsoftboundary 2.35	Widthofsoftboundary 7.52	Widthofsoftboundary 0.61
	DM_3 0 (72.0%)	DM_3 1 (100.0%)	DM_3 1 (66.2%)	DM_3 0 (100.0%)	DM_3 0 (100.0%)
	WidthbetweenboundarytreesW2m	WidthbetweenboundarytreesW2m	WidthbetweenboundarytreesW2m	WidthbetweenboundarytreesW2m	WidthbetweenboundarytreesW2m
	HeightofbuildingsboundaryH1m	HeightofbuildingsboundaryH1m	HeightofbuildingsboundaryH1m	HeightofbuildingsboundaryH1m	HeightofbuildingsboundaryH1m
	DL_1 0 (92.0%)	DL_1 0 (100.0%)	DL_1 0 (62.3%)	DL_1 1 (74.3%)	DL_1 0 (100.0%)
	DE_4 0 (66.0%)	DE_4 0 (100.0%)	DE_4 0 (81.8%)	DE_4 1 (81.1%)	DE_4 0 (100.0%)
	DL_5 0 (93.0%)	DL_5 0 (100.0%)	DL_5 1 (54.5%)	DL_5 0 (100.0%)	DL_5 0 (100.0%)
	DS_2 1 (60.0%)	DS_2 1 (68.4%)	DS_2 0 (100.0%)	DS_2 1 (73.0%)	DS_2 0 (100.0%)
	Id_7road_12sidewalk_53path_54stairs	Id_7road_12sidewalk_53path_54stairs	Id_7road_12sidewalk_53path_54stairs	Id_7road_12sidewalk_53path_54stairs	Id_7road_12sidewalk_53path_54stairs
	WidthbetweenbuildingsW1m	WidthbetweenbuildingsW1m	WidthbetweenbuildingsW1m	WidthbetweenbuildingsW1m	WidthbetweenbuildingsW1m
	HeightofboundarytreeH2m	HeightofboundarytreeH2m	HeightofboundarytreeH2m	HeightofboundarytreeH2m	HeightofboundarytreeH2m
	DE_9 0 (100.0%)	DE_9 0 (100.0%)	DE_9 0 (100.0%)	DE_9 0 (100.0%)	DE_9 0 (66.7%)
	Id_14earth 0.01	Id_14earth 0.09	Id_14earth 0.02	Id_14earth 0.07	Id_14earth 0.00
	DE_6 0 (70.0%)	DE_6 1 (50.5%)	DE_6 0 (100.0%)	DE_6 0 (100.0%)	DE_6 0 (100.0%)
	Id_3sky 0.15	Id_3sky 0.14	Id_3sky 0.09	Id_3sky 0.04	Id_3sky 0.20
	W1H1 3.04	W1H1 0.33	W1H1 0.72	W1H1 1.75	W1H1 2.27
	DM_4 0 (58.0%)	DM_4 0 (100.0%)	DM_4 0 (83.1%)	DM_4 0 (100.0%)	DM_4 0 (100.0%)
	DM_5 0 (100.0%)	DM_5 0 (100.0%)	DM_5 0 (96.1%)	DM_5 0 (100.0%)	DM_5 0 (72.5%)
	Id_1wall2building_26house	Id_1wall2building_26house	Id_1wall2building_26house	Id_1wall2building_26house	Id_1wall2building_26house
	Id_5tree_10grass_18plant	Id_5tree_10grass_18plant	Id_5tree_10grass_18plant	Id_5tree_10grass_18plant	Id_5tree_10grass_18plant
	DL_4 0 (87.0%)	DL_4 0 (100.0%)	DL_4 0 (100.0%)	DL_4 0 (74.3%)	DL_4 0 (100.0%)
	DE_7 0 (90.0%)	DE_7 0 (100.0%)	DE_7 0 (75.3%)	DE_7 0 (100.0%)	DE_7 0 (100.0%)
	DE_5 0 (74.0%)	DE_5 0 (50.5%)	DE_5 0 (70.1%)	DE_5 0 (85.1%)	DE_5 0 (100.0%)
	DE_3 0 (100.0%)	DE_3 0 (100.0%)	DE_3 0 (85.7%)	DE_3 0 (100.0%)	DE_3 0 (100.0%)
	DE_1 0 (100.0%)	DE_1 0 (100.0%)	DE_1 0 (87.0%)	DE_1 0 (100.0%)	DE_1 0 (100.0%)
	DM_1 0 (100.0%)	DM_1 0 (100.0%)	DM_1 0 (87.0%)	DM_1 0 (100.0%)	DM_1 0 (100.0%)
	Id_88streetlight 0.00	Id_88streetlight 0.00	Id_88streetlight 0.00	Id_88streetlight 0.00	Id_88streetlight 0.00
	Id_22water_61river 0.00	Id_22water_61river 0.02	Id_22water_61river 0.00	Id_22water_61river 0.03	Id_22water_61river 0.00
	Id_33fence 0.00	Id_33fence 0.01	Id_33fence 0.00	Id_33fence 0.01	Id_33fence 0.01
	DS_3 0 (68.0%)	DS_3 0 (78.9%)	DS_3 0 (100.0%)	DS_3 0 (73.0%)	DS_3 0 (66.7%)
	Id_35rock 0.00	Id_35rock 0.01	Id_35rock 0.00	Id_35rock 0.01	Id_35rock 0.00
	Id_21car 0.01	Id_21car 0.00	Id_21car 0.01	Id_21car 0.00	Id_21car 0.01
	WidthbetweenpavementsboundariesW3m	WidthbetweenpavementsboundariesW3m	WidthbetweenpavementsboundariesW3m	WidthbetweenpavementsboundariesW3m	WidthbetweenpavementsboundariesW3m
	DE_2 0 (100.0%)	DE_2 0 (100.0%)	DE_2 0 (100.0%)	DE_2 0 (95.9%)	DE_2 0 (100.0%)

(2) Distribution of features in five clusters—nominal variables

(3) clustering results of each sample (partial screenshots of the results are shown in the blue box because there are many sample features, all results cannot be displayed here).

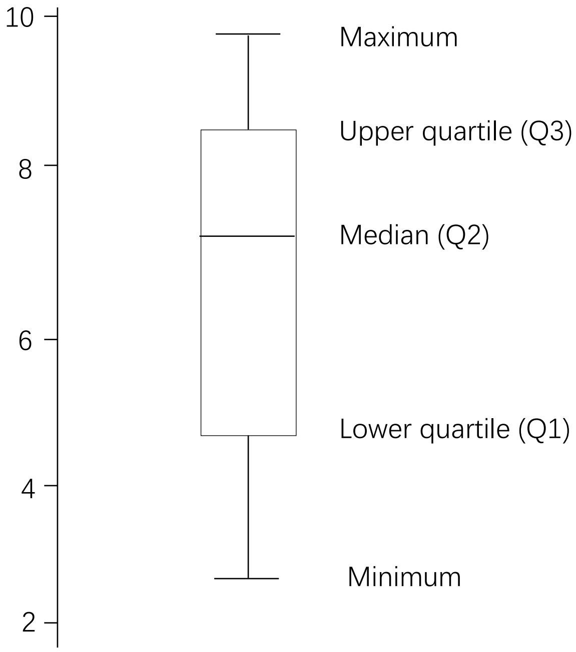


Sections	Photofile	ValenceB	ValenceT	Id_1wall12b	Id_3sky	Id_5tree_10	Id_7road	Id_14ear	Id_21car	Id_22wat	Id_33fen	Id_35roc	Id_88str	Widthbet	Widthbet
Section1	jaist (2	1.00	1.00	0.09	0.27	0.23	0.39	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (4	1.00	1.00	0.01	0.19	0.50	0.28	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (6	1.00	1.00	0.03	0.19	0.50	0.26	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (8	1.00	1.00	0.04	0.19	0.46	0.31	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (1	1.00	1.00	0.08	0.34	0.25	0.32	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (1	1.00	1.00	0.06	0.26	0.33	0.28	0.01	0.00	0.00	0.01	0.00	0.00	53.87	14.50
	jaist (1	1.00	1.00	0.07	0.27	0.28	0.24	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (1	1.00	1.00	0.10	0.26	0.31	0.27	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (1	1.00	1.00	0.15	0.25	0.22	0.33	0.00	0.00	0.00	0.00	0.00	0.00	53.87	14.50
	jaist (1	1.00	1.00	0.18	0.27	0.22	0.18	0.00	0.00	0.02	0.00	0.00	0.00	53.87	14.50
	jaist (2	1.00	1.00	0.34	0.14	0.21	0.19	0.00	0.00	0.00	0.01	0.00	0.00	53.87	14.50
Section2	jaist (2	1.00	1.00	0.00	0.03	0.72	0.13	0.00	0.00	0.00	0.00	0.00	0.00	41.87	5.50
	jaist (2	1.00	1.00	0.02	0.01	0.80	0.15	0.00	0.00	0.01	0.00	0.00	0.00	41.87	5.50
	jaist (2	1.00	1.00	0.00	0.04	0.80	0.11	0.01	0.00	0.01	0.01	0.00	0.00	41.87	5.50
	jaist (2	1.00	1.00	0.02	0.03	0.72	0.22	0.00	0.00	0.00	0.00	0.00	0.00	41.87	5.50
	jaist (3	1.00	1.00	0.14	0.04	0.58	0.22	0.00	0.00	0.00	0.00	0.00	0.00	41.87	5.50
	jaist (3	1.00	1.00	0.30	0.01	0.40	0.16	0.00	0.00	0.01	0.01	0.00	0.00	41.87	5.50
	jaist (3	1.00	1.00	0.07	0.05	0.68	0.13	0.03	0.00	0.00	0.01	0.00	0.00	41.87	5.50
	jaist (3	1.00	1.00	0.04	0.12	0.66	0.15	0.00	0.00	0.00	0.01	0.00	0.00	41.87	5.50
Section3	jaist (3	1.00	0.00	0.41	0.20	0.06	0.28	0.00	0.00	0.00	0.01	0.00	0.00	36.00	50.00
	jaist (3	1.00	0.00	0.29	0.20	0.07	0.26	0.03	0.00	0.00	0.01	0.00	0.00	36.00	50.00
	jaist (4	1.00	0.00	0.16	0.30	0.23	0.25	0.00	0.00	0.00	0.01	0.00	0.00	36.00	50.00
	jaist (4	1.00	0.00	0.13	0.33	0.21	0.26	0.00	0.00	0.00	0.01	0.00	0.00	36.00	50.00
	jaist (4	1.00	0.00	0.05	0.45	0.23	0.20	0.00	0.00	0.00	0.01	0.00	0.00	36.00	50.00
	jaist (4	1.00	0.00	0.00	0.30	0.29	0.33	0.01	0.00	0.00	0.00	0.00	0.00	36.00	50.00
	jaist (4	1.00	0.00	0.00	0.41	0.30	0.25	0.00	0.00	0.00	0.00	0.00	0.00	36.00	50.00
	jaist (5	1.00	0.00	0.07	0.39	0.29	0.23	0.00	0.00	0.00	0.00	0.00	0.00	36.00	50.00
	jaist (5	1.00	0.00	0.09	0.38	0.16	0.33	0.00	0.00	0.00	0.02	0.00	0.00	36.00	50.00
	jaist (5	1.00	0.00	0.21	0.27	0.22	0.28	0.00	0.00	0.00	0.01	0.00	0.00	36.00	50.00
	jaist (5	1.00	0.00	0.23	0.26	0.14	0.30	0.00	0.00	0.00	0.00	0.00	0.00	36.00	50.00
	jaist (5	1.00	0.00	0.28	0.30	0.07	0.34	0.00	0.00	0.00	0.00	0.00	0.00	36.00	50.00
	jaist (6	1.00	0.00	0.33	0.12	0.18	0.13	0.10	0.01	0.00	0.01	0.00	0.00	36.00	50.00

....

DE_5	DE_6	DE_7	DE_8	DE_9	TSC_6041_0.5OutlierRemoved
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
1.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	0.00	0.00	4.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00
0.00	0.00	0.00	1.00	0.00	2.00

Appendix G: Example of explaining the quartile values of the box diagram.



Appendix H: Normal distribution testing results of continuous variable.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Id_1wall-2building_26house	.181	437	.000	.845	437	.000
Id_3sky	.111	437	.000	.925	437	.000
Id_5tree_10grass_18plant	.105	437	.000	.933	437	.000
Id_7road_12sidewalk_53path_54stairs	.052	437	.007	.975	437	.000
Id_14earth	.284	437	.000	.641	437	.000
Id_21car	.377	437	.000	.349	437	.000
Id_22water_61river	.377	437	.000	.340	437	.000
Id_33fence	.307	437	.000	.573	437	.000
Id_35rock	.419	437	.000	.301	437	.000
Id_88streetlight	.374	437	.000	.371	437	.000
Width between buildings (W1)(m)	.131	437	.000	.874	437	.000
Width between boundary trees (W2)(m)	.161	437	.000	.875	437	.000
Width between pavement's boundaries	.293	437	.000	.581	437	.000
Height of building's boundary(H1)(m)	.109	437	.000	.967	437	.000
Height of boundary tree(H2)(m)	.132	437	.000	.957	437	.000
W1/H1	.159	437	.000	.789	437	.000
W2/H2	.175	437	.000	.789	437	.000
Width of soft boundary	.337	437	.000	.430	437	.000
Percentage of main space length	.150	437	.000	.921	437	.000
Public\ Private	.192	437	.000	.887	437	.000
Natural\Artificial	.157	437	.000	.914	437	.000
Modern\ Historical	.187	437	.000	.840	437	.000
Open\ Enclosure	.098	437	.000	.946	437	.000
Diversity\ Monotonous	.131	437	.000	.933	437	.000
Easy to identity\ Uneasy to identity	.125	437	.000	.917	437	.000
Green-rich\ Insufficient green	.142	437	.000	.891	437	.000
Unique\Common	.106	437	.000	.946	437	.000
Beautiful\ Ugly	.130	437	.000	.931	437	.000
Meaningful\ Meaningless	.114	437	.000	.935	437	.000
Artistic\ Inartistic	.098	437	.000	.964	437	.000
Continuous space\ Interrupted space	.180	437	.000	.908	437	.000

a. Lilliefors Significance Correction.