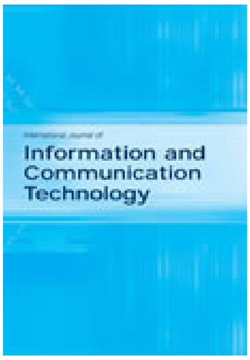


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Landscape perception analysis and design optimisation of urban parks supported by multimodal data

Yifei Zhang

DOI: 10.1504/IJICT.2025.10071989

Article History:

Received:	14 May 2025
Last revised:	25 May 2025
Accepted:	25 May 2025
Published online:	16 July 2025

Landscape perception analysis and design optimisation of urban parks supported by multimodal data

Yifei Zhang

Department of Media and Design,
GuiZhou Vocational Technology,
College of Electronics and Information,
Kaili 556000, China
Email: zhang815474413@163.com

Abstract: City parks are increasingly crucial in improving inhabitants' quality of life as well as other factors given the fast worldwide urbanisation. But as urban population density rises, the classic landscape design of urban parks has proved challenging to satisfy the ever varied needs of people. This work presents a multimodal data fusion-based landscape perception analysis and design optimisation method for urban parks, so building a comprehensive landscape perception model that can effectively capture the multidimensional traits of park environments and their effect on visitor experience, so solving these problems. Both tests reveal that multimodal data fusion greatly increases the comprehensiveness and accuracy of landscape perception. Although this study has achieved great progress, there are still certain constraints including the restricted capacity to react to dynamic data in real time; so, future studies will concentrate on addressing these issues and optimising the research approach.

Keywords: multimodal data fusion; urban park; landscape perception; intelligent optimisation.

Reference to this paper should be made as follows: Zhang, Y. (2025) 'Landscape perception analysis and design optimisation of urban parks supported by multimodal data', *Int. J. Information and Communication Technology*, Vol. 26, No. 26, pp.33–50.

Biographical notes: Yifei Zhang obtained his PhD from the Kunsan National University in South Korea in 2025. He is currently a Lecturer at the GuiZhou Vocational Technology College of Electronics and Information. His research interests include landscape design and multimodal data fusion.

1 Introduction

1.1 Research background and significance

Urban parks, as a crucial component of urban public space, help to improve the quality of life of people by means of acceleration of global urbanisation, therefore improving the urban environment and fostering social interaction (Hajzeri, 2021). But as urban population density rises, the classic landscape design of urban parks has been challenging to satisfy the increasingly varied needs of the public; problems including illogical functional layout and unequal arrangement of landscape elements usually arise in landscape design, so influencing the public's experience of use (Anguelovski et al., 2019). Consequently, in the field of urban planning and landscape design, how to scientifically and rationally implement the design and optimisation of urban park landscape becomes a crucial problem to be resolved.

Although traditional park landscape design mostly depends on the experience and intuition of designers, this approach is frequently difficult to achieve thorough examination and exact optimisation when confronting complicated environmental factors and diverse user needs. Thanks in great part to multimodal data, more and more urban parks are planned and optimised in recent years with the advance of information technology. Multimodal data is the several forms of data from several data sources, including images, sounds, environmental monitoring data, user comments from social media (Hangloo and Arora, 2022). Comprehensive analysis of these data helps one to have a more complete knowledge of all aspects of park environments and expose the link between landscape and user behaviour, so supporting design decisions.

The great utilisation of multimodal data enables landscape design to be maximised in a data-driven manner instead of depending just on human experience. By means of the analysis of park environment data, visitor behaviour data, visual and auditory data, etc. it is feasible to exactly pinpoint which landscape aspects people enjoy and which parts demand improvement (Hornecker et al., 2023). Simultaneously, these facts can enable designers to implement intelligent and tailored landscape design from the standpoint of user requirements. Thus, the study and optimisation of landscape perception in urban parks depending on multimodal data can not only enhance the accuracy and scientificity of design but also offer fresh ideas and approaches for the sustainable development of parks.

Based on multimodal data, this study intends to investigate the method of landscape perception analysis and design optimisation of urban parks, with an emphasis on how to extensively explore the intrinsic connection between urban park landscapes and user behaviours by data integration and model building, and so, propose scientific and reasonable landscape optimisation strategies. This study not only has great theoretical value and supports the development of data and intelligence in the field of landscape design but also has a wide range of social application prospects, which can provide data-driven decision support for the management and optimisation of urban parks.

1.2 Current status of related research and technology

As urbanisation accelerates constantly, perceptual analysis and urban park landscape optimisation has progressively become a prominent subject for study. With the advent of multimodal data capture technologies in recent years, researchers have started to investigate how to holistically examine urban park settings using several data sources (e.g., pictures, videos, environmental data, social media comments, etc.). For landscape design, multimodal data offers a more precise and complete basis for decision-making; it also makes data-driven design feasible.

Urban park landscape perception analysis makes extensive use of technology in computer vision and image processing. From images and videos of park environments, researchers are able to extract various landscape features, including plants, buildings, and routes, by employing techniques including image recognition and target identification, then analyse the impact of these elements on visitor attraction (Van Berkel et al., 2018). For automatic detection and categorisation of landscape elements, for instance, convolutional neural network (CNN) is a classic deep learning model extensively applied. After extracting landscape features from CNN analysis of park landscape images, then the study evaluated the appeal of these characteristics to tourists. Researchers have thus started to investigate how to incorporate various kinds of multimodal data to increase the accuracy of perceptual analysis as the processing and analysis of picture data still faces the difficulties of environmental complexity and great data volume.

Apart from picture information, environmental data is also crucial for landscape perception study. Sensor technology has advanced to allow environmental data in parks to be continuously recorded. Many studies have examined, using environmental data and visitor behaviour data, the influence of environmental elements on tourist experience (Wang et al., 2019). For instance, some research using machine learning techniques including support vector machine (SVM) and decision tree has modelled the link between environmental elements and visitor length of stay. By means of these models, researchers can forecast the likely behaviours and preferences of guests under various environmental settings, therefore offering theoretical help for the optimisation of park landscapes. In this regard, decision tree is very helpful since it allows one to graphically show how different environmental factors affect tourists' decisions (Serrano López et al., 2019).

As a newly available data source, social media data has also been extensively applied in perceptual study of park landscapes in recent years. Using crawling photographs, comments, geographic locations, and other data on social media sites, researchers have examined tourists' emotional responses and behavioural paths towards park settings (Barros et al., 2020). Sentiment analysis methods, for instance, were used to evaluate social media text data and by means of sentiment analysis of visitor comments, researchers were able to pinpoint which aspects of the landscape were popular and which ones might use improvement. Furthermore, social media location data allows researchers to map visitors' activity heat maps, therefore exposing their park length of stay and travel routes. These statistics offer fresh concepts for besting park functional layout.

Regarding multimodal data fusion and landscape optimisation, new work has also shown several sophisticated machine learning and deep learning methods. An essential optimising tool, reinforcement learning (RL), has lately been used in the intelligent design of urban park environments. By means of RL, researchers may replicate the behavioural process of visitors in the park, modify the landscape design in response to tourist comments, and maximise the division of the functional regions of the park and the arrangement of the landscape elements. For instance, the researchers continuously changes the design using the reward and punishment mechanism to eventually get the ideal landscape layout scheme and employs RL algorithms to replicate the visitors' visiting behaviour under various landscape design schemes (Ali et al., 2023). This approach avoids the restrictions of the hand design process by automatically discovering the optimal option among several design ideas.

Apart from RL, generative adversarial networks (GAN) in deep learning have also been experimented upon in landscape design optimisation. To create customised landscape layouts, the researchers continuously improved the produced designs depending on user feedback and developed new landscape design patterns using GAN (Lu et al., 2024). This method offers a fresh perspective on landscape design that might produce rather original and individualised solutions backed by a lot of data.

Furthermore, applied in landscape analysis and optimisation are ensembles of learning and clustering methods. Combining several weak learners helps integration learning techniques like Random Forest to increase the accuracy of categorisation and evaluation of landscape features. Conversely, clustering techniques can be applied to group visitor activity, pinpoint the demands of various kinds of guests, and so targetably improve park environments depending on these needs.

As technology develops, more and more algorithms and techniques are being used to the perceptual analysis and design optimisation of metropolitan park environments. From image processing and environmental data analysis to social media data mining to the application of machine learning and deep learning algorithms, these technologies bring new viewpoints and means for landscape design. Though much of the current research has advanced, how to effectively combine multi-source data and create more accurate models to handle complicated multimodal data remain difficulties in present work. Future study in this field so still revolves around more effective data fusion methods and optimisation algorithms.

1.3 Innovations and contributions of the study

This work presents a thorough investigation in the subject of landscape perception research and design optimisation of urban parks and suggests a new optimisation method depending on multimodal data fusion. The innovations and contributions of this study mostly show themselves in the following features when compared with previous studies:

1. A framework for analysing the landscape perception of urban parks based on multimodal data fusion is proposed: an integrated perception model is created by combining multimodal data including photographs, environmental monitoring data, user behaviour data and social media comments, therefore enabling a more complete and accurate analysis of the multidimensional qualities of the park landscape. This method is innovative in that it can evaluate the park landscape from several angles concurrently by means of several data sources, therefore enhancing the accuracy and applicability of the analytical findings.

2. A data-driven intelligent optimisation method is introduced in landscape design optimisation: this approach allows the automatic analysis and identification of the main components of the park landscape, thereby optimising the arrangement of the landscape features and the separation of functional regions based on visitor preferences and behaviour pattern. Particularly by means of RL algorithms, the study is able to dynamically adapt and optimise in several design situations, so enhancing the accuracy and efficiency of landscape design.

3. Innovative attempts are made in data fusion technology: in the process of multimodal data fusion, this work uses advanced data pre-treatment and feature fusion techniques to fully play the complementary character of every data source by means of efficient processing and merging of various kinds of data. This data fusion approach not only enhances the predictive capacity of the model but also offers better information support for later landscape optimisation.

4. The application value of the study is of great significance: the study offers fresh viewpoints and technical tools for the design and management of urban parks by suggesting a multimodal data fusion-based landscape perception analysis and design

optimisation method for them. The findings of the study not only offer fresh theoretical support for related research in academia but also offer tools and approaches with pragmatic relevance for the actual running and administration of urban parks.

In essence, this work presents a thorough analysis framework for multimodal data fusion and solves accuracy and personalisation issues in conventional landscape design by means of an intelligent optimisation approach, so advancing the research development of data-driven park landscape design. These developments and efforts not only improve the theoretical framework in the field of landscape design but also give scientific planning and sustainable development of urban parks efficient technical support.

2 Multimodal data and urban park landscape perception

2.1 Definition and application of multimodal data

Multimodal data is heterogeneous collection of several kinds from several sources or sensors. Images, videos, text, music, sensor data, etc. reflecting many facets of the same object or phenomenon in several dimensions, can be included in these data. Although many data kinds in multimodal data have diverse structures and formats, they are essentially linked to one another. Effective integration and analysis of this data helps to clarify and characterise the target object or phenomena, therefore offering more complete information support for next decisions.

Application of multimodal data is quite important in the perceptual study and design optimisation of urban park environments. Park landscape encompasses visitors' visual, aural, emotional, and other components of experience in addition to being a physical environment (Zheng et al., 2020). Consequently, a single data source usually cannot adequately represent the features of the terrain and the requirements of visitors. Image data, for instance, can show the terrain, environmental sensor data can represent the temperature, humidity, and other environmental conditions in the park, and social media data offers emotive comments and visitor behaviour paths (Rokhsaritalemi et al., 2023). Combining these disparate data enables a thorough evaluation of the components of the park landscape and their influence on the tourist experience.

In related disciplines, the combination of multimodal data is now a fundamental technology. Data fusion allows correlations between several modalities to improve the performance of analytical models to be created (Gao et al., 2020). Common methods in multimodal data fusion are early, late, and intermediate fusion. Early fusion is the combining of data from many modalities at the data input stage for unified processing; late fusion is the integration of results following the analysis results of each data modality are available; and mid-term fusion is the combination of data from many modalities at the feature extraction stage. Usually, the particular application situation and data properties determine the fusion method to be used.

Assuming image data X_{image} and environmental data X_{env}, the fusion of the two data sources can be stated using urban park landscape optimisation as an example:

$$X_{\text{fused}} = f\left(X_{\text{image}}, X_{\text{env}}\right) \tag{1}$$

where X_{fused} is the fused multimodal data; $f(\cdot)$ is the fusion function, either a straightforward weighted average or a more advanced deep learning network. Particularly in the context of image and sensor data fusion, deep learning models such CNNs and recurrent neural networks (RNNs) can efficiently capture the nonlinear relationships between several data modalities, so becoming one of the mainstream techniques in practical applications.

Practically, the combination of multimodal data can not only enhance the general impression of the landscape but also enable the development of tailored design optimisation solutions. Other studies understand tourists' preferences by combining social media data and image data, so providing data support for landscape design; some studies analyse how environmental factors affect tourists' emotional fluctuations by combining tourists' emotional feedback data with environmental data, so guiding the comfort enhancing of landscapes (Lin and Yang, 2024).

Consequently, the use of multimodal data can significantly raise the degree of customisation and precision in the urban park landscape planning. Future landscape design will depend much on how effectively to combine this disparate data and enhance the fusion effect.

2.2 *Urban park landscape perception methods*

Urban Park landscape perception is the process by which different technical approaches enable one to perceive and analyse environmental conditions, visitor behaviour, and landscape aspects in parks. Parks are becoming more and more significant public spaces in cities as urbanisation accelerates; so, the best design of their landscapes is very vital for enhancing the quality of life of the people and supporting sustainable urban growth. Landscape perception encompasses the integration of multi-dimensional information including visitors' behavioural patterns, environmental perception and social feedback in addition to a basic account of the visual features of the landscape. Consequently, computer vision, sensor technology, and social media analysis are among the several data collecting and analysis tools used in landscape perception approaches for urban parks.

Image and video analysis methods have been extensively applied for automatic detection and classification of landscape features in computer vision. By means of image acquisition and processing, landscape features in parks can be automatically identified and relevant feature information extracted. CNN analysis of park photos, for instance, can produce automatic landscape element classification and annotations (Lee and Son, 2022). Assuming X_{image} as the image data, which comprises park landscape elements, the CNN may extract the feature information via the convolution process, stated as:

$$F_{\text{image}} = \text{CNN}\left(X_{\text{image}}\right) \tag{2}$$

where F_{image} indicates the feature vector obtained by CNN, which comprises the visual information of landscape elements including the area of the green space and the landscape distribution. By means of these elements, one can investigate further the park's landscape design, visitor activity paths, and landscape's utility.

Moreover, landscape perception depends much on environmental sensor technologies. By means of the sensors placed in the park, the environmental conditions can be continuously monitored and hence directly influences the visitors' experience as well as the ecological purpose of the park. Data on temperature and humidity, for instance, might mirror the degree of comfort in particular park sections, which influences visitors' activity time and preferred stay (Li et al., 2018). The following formula allows one to represent the environmental state in the park at a given instant assuming that the data acquired by the environmental sensor is X_{env}:

$$X_{\text{env}} = [T, H, L, N] \tag{3}$$

where T stands for temperature; H stands for humidity; L stands for light intensity; N stands for noise level. Different landscape areas' comfort degrees as well as their effects on visitor behaviour can be evaluated using environmental data. By means of data analysis, researchers can pinpoint which environmental elements significantly influence visitor preferences and hence, maximise the park's functional zoning (Liang and Li, 2023).

Combining social media data, location data, and behavioural data helps visitor behaviour analysis to provide a more exact knowledge of needs and preferences. Pictures and comments as well as visitors' social media entries help to capture their emotional views and tastes toward particular landscape areas. Furthermore, using GPS position data, one can monitor visitors' activity paths and examine their park behaviour. By means of clustering algorithms, for instance, analysing behavioural data of tourists, it is feasible to determine the several forms of behaviour, including recreational, sporty, sociable, etc. and create tailored landscape areas depending on these forms.

In landscape perception, sentiment analysis methods are also somewhat extensively applied. Analysing visitors' social media comments and sentiment feedback helps academics to grasp the appeal of various landscape components to them. To find the landscape sections that visitors prefer, a sentiment analysis algorithm, for instance, can sort comments into positive, negative, and neutral categories. Using the sentiment analysis model assuming X_{social} as the comment data, one can find the sentiment score with the following formula:

$$S_{\text{social}} = SentimentAnalysis\left(X_{\text{social}}\right) \tag{4}$$

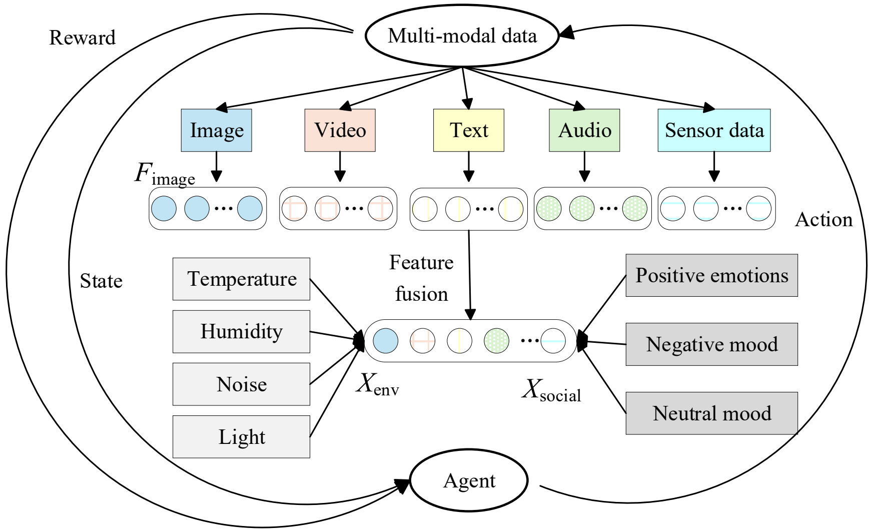
Among them, S_{social} indicates the sentiment score of social media comments, therefore reflecting the general opinion of guests on the park environment. This allows landscape designers to make deliberate changes to the park environment depending on comments from visitors therefore improving the whole experience of guests.

Methods of urban park landscape perception address image processing, environmental sensing, social media analysis, etc. Combining and assessing these approaches helps one grasp the park scene, visitor needs, and data for landscape design and park optimisation. These approaches offer fresh concepts for the building of sustainable urban parks as well as a theoretical framework for smart and customised landscape optimisation.

3 Design optimisation methodology

A landscape perception study and design optimisation technique for urban parks is proposed using multimodal data fusion. Figure 1 presents a complete landscape perception model by means of the combination of picture, environmental monitoring, user activity, and social media data. The model might more precisely evaluate the multidimensional qualities of park settings, therefore supporting data for improvement of the landscape.

Figure 1 Urban park landscape perception design method (see online version for colours)



Specifically, the design optimisation methodology of this study consists of the following main components:

3.1 Multimodal data fusion and landscape perception analysis

Improving analysis accuracy in urban park landscape perception study depends on the integration of information from many data sources. Multimodal data fusion's main goal is to combine several kinds of data, (e.g., pictures, environmental monitoring data, user behaviour data, and social media comments) into a single model to fairly depict the multidimensional aspects of park environments. Combining these data helps the study to fully capture the visual, environmental, and social elements of park settings, hence enhancing the capacity for analysis of park environments.

First of all, image data can show the landscape's visual characteristics including the colour, form, and texture of the landscape components (Karasov et al., 2021). CNN helps one to extract the features of the landscape image so obtaining its high-dimensional feature representation. Particularly, the image feature vector x_i reflects the visual information of every element in the park landscape and is obtained from it. One may get the picture feature vector's representation by means of the following equation:

$$x_i = f_{cnn}\left(I_i\right) \tag{5}$$

where I_i is the picture of the park scenery; f_{cnn} stands for CNN model.

Second, the environmental monitoring data which include temperature, humidity, light intensity, etc. support the ecological aspects of the terrain. Denoted as e_i, the environmental data feature vector is obtained assuming n sensors from environmental data collecting.

$$e_i = [T_i, H_i, L_i] \tag{6}$$

where T_i shows temperature; H_i shows humidity; L_i shows light intensity; all of these together reflect the environmental qualities of the park.

In multimodal data fusion, then, many kinds of data have to be combined in the same feature space. Weighted average or deep learning methodology allows us to obtain the fused feature vector z_i assuming knowledge from m data sources (including photos, environmental data and behavioural data, etc.):

$$z_i = \sum_{k=1}^{m} \alpha_k f_k \tag{7}$$

Reflecting the weight of every data source in the fusion process, f_k is the feature vector of the k^{th} data source while α_k is the associated weight coefficient. By means of weighted summing, the information from several data sources is combined into a single feature vector, therefore enabling efficient fusion of multimodal data.

By means of this multimodal data fusion method, the research is able to capture the features of park landscapes in several dimensions, including visual, environmental, and user behaviours, so offering complete data support for landscape perception analysis.

3.2 Application of intelligent optimisation methods and RL

The implementation of intelligent optimisation approaches offers great help for the layout of landscape elements and the division of functional sections in the optimisation of urban park landscape design. These optimisation techniques can automatically examine and pinpoint important components in the park environment and adjust the design plan in line with visitor tastes and behaviour. Specifically, by means of a feedback mechanism in an always changing environment, RL, as an adaptive optimisation method, may constantly improve the landscape design. RL allows complicated urban park design to iteratively optimise the design scheme from several dimensions and adapt it depending on the feedback information, so obtaining accurate and effective design optimisation.

First in the intelligent optimisation strategy is definition of states, actions, and reward systems. While the present condition of the design of the park design can be seen as a result, every element arrangement of the design can be considered as an action in the process of optimisation of urban park landscape design. Every change in the design is assessed using a reward function, generally connected to elements like environmental comfort and visitor pleasure (Mfon, 2023). For instance, real-time adjustment of landscape design solutions depending on visitor behaviour data, environmental input, etc. helps to maximise the allocation of the functional area in the park.

The fundamental concept of RL is to progressively optimise the decision-making strategy by means of the reward and punishment mechanism between the intelligent body and the surroundings. It is specifically believed that the state of every design scheme may be expressed as s_t, where t signifies the time step. By choosing the best design scheme, RL aims to maximise the long-term reward R and may be stated as follows:

$$R = \sum_{t=0}^{T} \gamma^t r_t \tag{8}$$

where T is the whole period of optimisation; r_t is the immediate reward at time step t; γ is a discount factor deciding the weight of the future reward.

The decision-making process of RL intelligences must be modelled by a state transfer function if one wants to maximise in an environment of changing design. Every state and action mix in park landscape design produces a new state. Changing the park's layout and the location of landscape features, for instance, can influence visitor behaviour, therefore inducing a state change. One may depict the state transfer by the following equation:

$$s_{t+1} = f(s_t, a_t) \tag{9}$$

where s_t represents the current state; a_t indicates the currently selected action; f stands for the function of state transfer.

The Bellman equation drives the process of changing the Q-value in RL. Long-term payoff of a particular activity in a given condition reflected by the Q-value and the update formula for the Q-value is:

$$Q(s_t, a_t) = (1-\alpha)Q(s_t, a_t) + \alpha\left[r_t + \gamma \max_{a'} Q(s_{t+1}, a')\right] \tag{10}$$

where α is the learning rate; r_t is the instantaneous reward; γ is the discount factor; $\max_{a'} Q(s_{t+1}, a')$ shows the maximum Q value in the next state. This defines the step size of the Q value update.

The RL intelligences may maximise the design scheme in continual interaction and change the design in real time depending on the feedback of the park landscape by applying the above formula, so obtaining superior optimisation (Massaoudi et al., 2023). Furthermore, deep learning technology can be coupled with the RL optimisation technique to handle complicated and high-dimensional design space, so improving the optimising impact.

Finally, the use of intelligent optimisation techniques allows urban park landscape design to rapidly react to environmental changes and visitor needs, thereby dynamically adjusting. Park design may not only increase efficiency but also better satisfy the several wants of visitors by means of the feedback mechanism of RL, thereby attaining ideal design in a dynamic environment.

3.3 Data pre-processing and feature fusion techniques

Key measures to increase the accuracy of landscape perception analysis in multimodal data fusion include data pre-treatment and feature fusion methods. Usually with varying formats, sizes, and noise, urban park landscape research incorporates a range of data types which include images, sensor data, user behaviour data, and social media data. Data preparation aims to remove unnecessary information and mistakes by means of cleaning, denoising, normalisation, and normalisation to produce high-quality inputs for next investigations (Massaoudi et al., 2023). Image data, for instance, might need to be denoised and normalised in size; sensor data might have missing values that need to be filled in or smoothed out. Conversely, feature extraction which involves sentiment analysis, topic identification, and interaction frequency extracts salient features from many kinds of data, including colour, texture, and shape features in images, temperature, humidity, and noise in environmental data, and social media data.

Feature fusion methods are applied to efficiently merge data features from several sources following data preparation, so utilising the complementing character of every form of data. Based on their signal or relevance, weighted fusion techniques sometimes give each modality varying weights (Ma et al., 2019). The weighted fusion formula can be stated especially as:

$$x_{\text{fused}} = \sum_{i=1}^{N} w_i x_i \tag{11}$$

where x_i is the feature vector of the i^{th} data class; w_i is the weight of the modality; N is the total number of modalities; the resultant x_{fused} is the fused feature vector.

This work uses another weighted fusion technique, constantly adjusted depending on the contribution of every data modality to the final target, hence enhancing the effect of feature fusion. An evaluation function helps one to determine the value of every data modality; so, the weight is changed based on their contribution to the outcome. At this moment, the fusion process may be stated with the following equation:

$$x_{\text{fused}} = \sum_{i=1}^{N} \alpha_i \cdot f\left(x_i\right) \tag{12}$$

The dynamic weight of the data modality of class i, $f(x_i)$ indicates the feature vector following pre-processing or feature extraction of class i data; the resultant x_{fused} is the fused feature vector. This formula allows the dynamic weight α_i to be dynamically changed depending on the significance of the data modality thereby optimising the fusion effect.

By means of this deliberate feature fusion approach, the quality and comprehensiveness of landscape perception research can be improved by fairly aggregating information from many data sources, therefore offering more correct inputs for later landscape design optimisation.

3.4 Real-time adjustments and personalised landscaping

Real-time adaptation and customisation improve the user experience in landscape design by means of constant collecting and analysis of multimodal data (Gómez-Carmona et al., 2022). Using sensor data, social media comments, and location information, the system can monitor visitor activity in real time and examine park visitor trajectory. For instance, the system dynamically changes landscape design components, (e.g., adding or eliminating resting spaces, optimising pathways, etc.) to match visitor demand when the density of visitors in a given location is too high or a landscape element is used less often. Real-time adjustment depends on efficient multi-source data processing as well as dynamic design optimisation based on changes (Chen et al., 2021).

Analysing visitor group preferences helps one to create a personalised landscape. The system uses visitor behavioural data to determine the preferences of various visitor groups, e.g., younger visitors like sports facilities while senior visitors prefer quieter places. This information helps the system to optimise the arrangement of landscape features so improving visitor happiness and spatial economy.

This work uses an intelligent optimisation method coupled with RL for dynamic design optimisation to reach these modifications. With RL, the system may continuously optimise the design solution. The reward function $R(s, a)$ for instance is defined as a function of visitor happiness and design solutions:

$$R(s, a) = \gamma \cdot \text{Satisfaction}(s, a) - \text{Cost}(a) \tag{13}$$

where γ is the discount factor; a is the design modification; s is the current landscape state. The system learns the consequences of various actions, therefore progressively improving the design scheme.

$$\hat{d}_{t+1} = \alpha \cdot d_t + (1-\alpha) \cdot d_{t-1} \tag{14}$$

where $\hat{d}_{t+1}$ is the demand estimate for the next moment; d_t is the current demand; α is the smoothing factor; d_{t-1} is the historical demand. By means of this forecast, the system can forwardly arrange and maximise the landscape design to guarantee that it is fit for the variations in visitor demand.

4 Experimentation and evaluation

4.1 Dataset and experimental setup

The landscape perception of urban parks is analysed and optimised in this work using the OpenStreet Map (OSM) dataset. Rich geospatial information from the OSM dataset spans thorough data of parks, green areas, roads, buildings, and other aspects. These facts can help the spatial analysis, functional zoning, and optimisation design of metropolitan park environments really well. Table 1 presents the key contents of the OSM dataset utilised in this investigation in order to guarantee the efficiency of the experiment.

Table 1 OSM dataset information

Data type	*Description*	*Data items*
Park area data	Geographic information on park locations, boundaries, and area size	Park name, location coordinates, boundary box, area size
Facility data	Distribution of facilities within the park, such as benches, toilets, and trash bins	Facility type (benches, toilets, rest areas), location coordinates, quantity
Road and path data	Information on roads and pathways within and around the park	Road name, path type, start/end location, length
Natural landscape data	Information on natural elements within the park, such as greenery, water bodies, and plants	Landscape type (lakes, grass fields, flower beds), location coordinates, area size
Urban area data	Surrounding urban area information such as streets, business districts, and transportation facilities	Street name, commercial area location, transportation facilities
Population and movement data	Social media and sensor data to analyse visitor activities and preferences	Visitor locations, stay duration, activity preferences

Data collecting and pre-processing, data fusion, model training and optimisation constitute the experimental framework of this work. First, OSM API helped us to obtain the necessary spatial data of urban parks: park boundaries, roads, facilities, green areas and other information. To guarantee a consistent data format and eliminate erroneous data, the acquired data were cleaned and standardised. Furthermore, developed was a multimodal data fusion system by aggregating environmental monitoring data, (e.g., air quality, temperature, etc.) with social media (e.g., Twitter, Instagram.).

The fused data are then trained. Simultaneously, several design ideas were refined depending on RL to dynamically change the arrangement of functional zones to satisfy visitor wants and preferences and landscape components. By use of multimodal data, the study was able to evaluate the multidimensional aspects of park environments more fully and so enhance the accuracy and efficiency of landscape design.

At last, the experiment compares several design alternatives in several facets to evaluate their performance and assesses the optimisation results to give a scientific basis for landscape design and optimisation in urban parks.

4.2 Experimental results and analysis

Two studies were carried out in this work to validate the multimodal data fusion design optimisation strategy for urban park landscape perception. Experiment 1 quantitatively analyses the contribution of multimodal data sources to landscape perception in urban parks, so verifying the efficacy and practicality of multimodal data fusion by means of their percentage contribution to landscape perception. The results of the experiment are shown in Table 2 and seek to evaluate in landscape perception the relative value of image data, environmental data, social media data, visitor behaviour data, time series data, and geographic data.

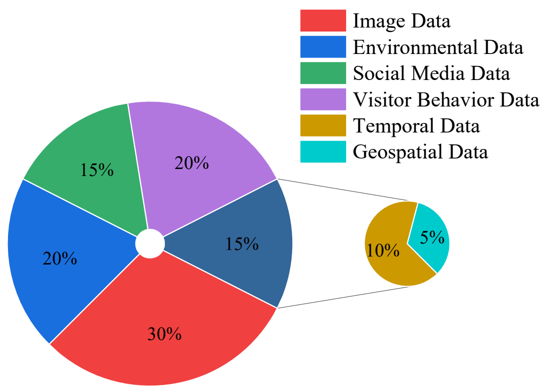
Table 2 Results of the contribution of multimodal data sources to landscape perception

Data source type	*Feature importance score*	*Contribution ratio (%)*	*Description*
Image data	0.30	30	Provides visual features of the landscape, such as green space area and path distribution.
Environmental data	0.20	20	Reflects climate conditions within the park, such as temperature, humidity, and light intensity.
Social media data	0.15	15	Provides visitor feedback and behavioural trajectories, such as sentiment analysis and activity paths.
Visitor behaviour data	0.20	20	Provides visitor activity trajectories and dwell time.
Temporal data	0.10	10	Analyses the impact of different seasons and time periods on landscape perception.
Geospatial data	0.05	5	Provides terrain, elevation, and slope information of the park.

CNN extracts the visual aspects of the landscape elements, including the area of green space and path distribution, from the OSM dataset. Sensors in the park gather environmental data like temperature, humidity, light intensity and noise level. Platforms such Twitter and Instagram provide the social media data, which comprises sentiment analysis findings, visitor comments, and activity paths. Furthermore, included into the trial were visitor behavioural data, derived from GPS positioning and camera observations in the park from activities trajectories and length of stay. Different seasons (spring, summer, autumn, winter) and time periods (weekdays, weekends) were investigated using time series data on the impression of the park scenery. On the other side, geospatial data which included topography, elevation, and park slope were derived from the OSM dataset.

Starting with pre-processing of every data source, including data cleaning, normalisation and feature extraction, the experimental processes were exacting. Then, a multimodal data fusion model is built to create a complete feature vector from the above six data sources. Using a feature importance evaluation approach, the experiment quantified the contribution of every data source to landscape perception. Comparatively analysing the model output with the manually labelled landscape features helped to confirm the model's accuracy at last. Further presentation of the experimental data is in Figure 2.

Figure 2 Feature importance scores for multimodal data (see online version for colours)



Reaching 30% and 20%, respectively, image data and visitor behaviour data account for a significant proportion of contribution in urban park landscape impression. Rich visual information from image data which are absolutely vital for evaluating the beauty and usefulness of the landscape is given by them. By means of activity trajectories and length of stay, visitor behaviour data on the other hand directly shows the real use and preference of visitors to various landscape areas. The great contribution ratio of these two datasets suggests that the most important determinant of landscape perception is actual visitor behaviour as well as visual components. Furthermore, shown by environmental data was a more noteworthy contribution of 20%. This implies that visitor experience and general impression of the environment are rather influenced by meteorological conditions, (e.g., temperature, humidity, and light intensity) within the park. Environmental elements not only influence visitor comfort but also might have an indirect impact on the visual effect and use function of the landscape.

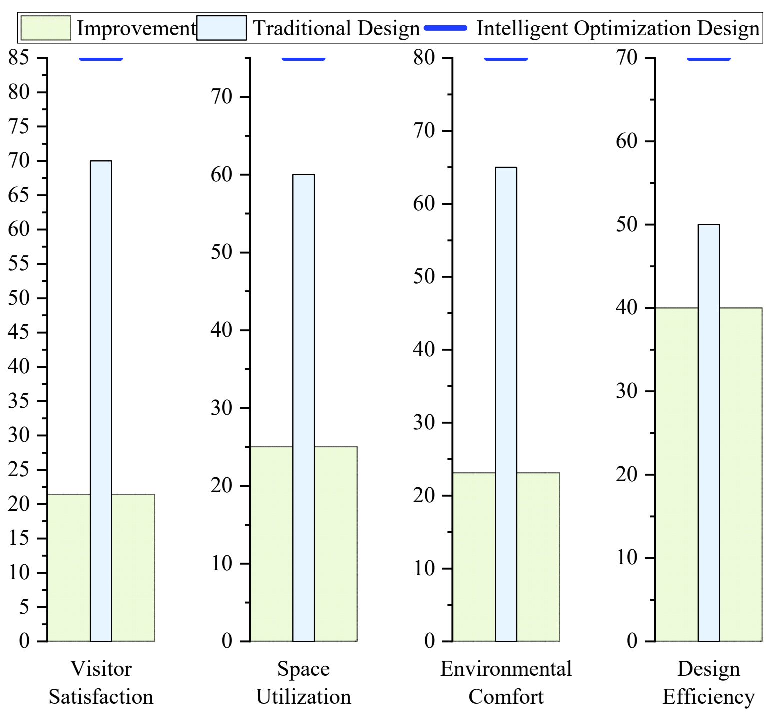
Social media data and time series data made 15% and 10% respectively contributions. By means of comments and sentiment analysis, social media data offer emotional feedback and subjective assessments of the environment. This information helps to grasp visitors' level of satisfaction with the landscape and recommendations for development. Conversely, time series data showed how various seasons and times affect landscape perception, which is crucial for long-term planning of the park and dynamic management.

Although geospatial data makes a quite small contribution at 5%, the information it offers on topography, elevation, and slope is nevertheless very helpful for knowing the general spatial structure and layout of the park. The multimodal data fusion technique offers rich data assistance for landscape design and optimisation, therefore reflecting the multifaceted qualities of urban park environments.

Using multimodal data gathered and processed in experiment 1, the second data preparation session covers image data, environmental data, social media data, visitor behaviour data, time series data, and geospatial data. The chosen research item for the experiment was a particular park, whose current landscape design scheme was created by a professional landscape architect drawing on conventional knowledge. First designed in the trial stages using conventional design principles, the landscape design concept of the park covered the arrangement of landscape elements and the division of functional zones. Then, utilising multimodal data fusion, an intelligent optimisation technique was presented to dynamically modify the division of functional areas and the arrangement of landscape features using the RL algorithm. By gathering environmental feedback data and real-time visitor behaviour data, this approach constantly improves the design scheme. The experiment runs the two design strategies in a simulated environment and gathers environmental feedback data, (e.g., temperature, humidity) and visitor behaviour data (e.g., dwell time, activity trajectory). At last, the performance of the two design approaches in terms of visitor satisfaction, space use and environmental comfort was evaluated holistically.

The experimental results are clearly presented through Figure 3.

Figure 3 Verification of the application effect in landscape design



In many important criteria, the intelligent optimisation approach based on multimodal data fusion clearly beats the conventional design solution according the evaluation results. With regard to visitor satisfaction from 70% to 85%, space use from 60% to 75%, environmental comfort from 65% to 80%, and design efficiency from 50% to 70%, the smart optimisation solution raised These developments not only greatly improve the general functioning of the park but also give urban designers and landscape architects a more scientific and effective design tool.

Overall, the outcomes of experiment 2 confirmed the efficiency and applicability of the intelligent optimisation approach grounded on multimodal data fusion in urban park landscape design. Introducing multimodal data and intelligent optimisation algorithms will not only increase the accuracy and scientificity of landscape design but also help to better satisfy the varied needs of visitors and offer fresh viewpoints and technical tools for the management and planning of urban parks.

5 Summary and outlook

5.1 Summary of the study

Aiming to improve the science and accuracy of landscape design by combining several data sources, this work presents a multimodal data fusion-based landscape perception analysis and design optimisation approach for urban parks.

This study generates a comprehensive landscape perception model able to capture the features of park landscapes and their effects on visitors' experiences from many dimensions by combining image data, environmental data, social media data, visitor behaviour data, time series data and geospatial data. With picture data and visitor behaviour data mostly contributing most importantly to landscape perception, the experimental results reveal that multimodal data fusion greatly increases the accuracy and comprehensiveness of landscape perception. Furthermore, performing well in landscape design optimisation, particularly in improving visitor happiness, space use, environmental comfort and design efficiency, is the intelligent optimisation technique based on multimodal data fusion.

These findings not only offer fresh ideas and approaches for the sustainable development of cities but also strong support for the development of data and intelligence in the field of landscape design, so promoting the shift from experience-driven to data-driven landscape design of urban parks, and provide new means for the planning and management of urban parks.

5.2 Research limitations and future work

There are still certain restrictions even if this work has achieved great advancement in multimodal data fusion and intelligent optimisation approaches. First of all, the present multimodal data fusion model has limited capacity to react in real time to dynamic data, (e.g., real-time visitor activity and environmental changes) and mostly depends on stationary data sources, such photos and environmental monitoring data. Second, even if the intelligent optimisation method shows good performance in tests, the generalisation capacity and adaptability of the model still need to be further confirmed in practical uses, particularly in parks of various sizes and kinds. Furthermore, the present work mostly addresses the optimisation of landscape design and the evaluation of long-term ecological effects and socio-cultural values of landscapes has not yet been thoroughly discussed.

Future activity can be conducted in the following spheres:

1 Dynamic data fusion and real-time response: the creation of multimodal fusion models able to analyse dynamic data in real-time to enhance the real-time responsiveness of landscape sensing systems would be the main emphasis of next studies. More sophisticated sensor technology and real-time data processing algorithms will be part of this in order to better record environmental changes and visitor activity.

2 Model generalisation and adaptation validation: future studies will evaluate the generalisation capacity and flexibility of the model by means of experiments in more kinds of parks so ensuring the efficiency of the intelligent optimisation technique in various surroundings. This will enable the algorithm to be even more optimised so that it may respond to various demands for landscape design.

3 Ecological and socio-cultural value assessment: future studies will cover the evaluation of socio-cultural values of landscapes and long-term ecological consequences. Combining multimodal data and incorporating ecological and social evaluation indicators will help to create a more complete landscape assessment model supporting sustainable landscape design and management.

By means of these future research areas, our work will enhance the multimodal data fusion and intelligent optimisation techniques to offer more complete and scientific technical assistance for urban park landscape design and so promote the sustainable growth of urban parks.

Declarations

All authors declare that they have no conflicts of interest.

References

Ali, M., Naeem, F., Kaddoum, G. and Hossain, E. (2023) 'Metaverse communications, networking, security, and applications: research issues, state-of-the-art, and future directions', *IEEE Communications Surveys & Tutorials*, Vol. 26, No. 2, pp.1238–1278.

Anguelovski, I., Irazábal-Zurita, C. and Connolly, J.J. (2019) 'Grabbed urban landscapes: socio-spatial tensions in green infrastructure planning in Medellín', *International Journal of Urban and Regional Research*, Vol. 43, No. 1, pp.133–156.

Barros, C., Moya-Gómez, B. and Gutiérrez, J. (2020) 'Using geotagged photographs and GPS tracks from social networks to analyse visitor behaviour in national parks', *Current Issues in Tourism*, Vol. 23, No. 10, pp.1291–1310.

Chen, G., Zhang, X., Wang, C., Zhang, Y. and Hao, S. (2021) 'Research on flexible control strategy of controllable large industrial loads based on multi-source data fusion of internet of things', *IEEE Access*, Vol. 9, pp.117358–117377.

Gao, J., Li, P., Chen, Z. and Zhang, J. (2020) 'A survey on deep learning for multimodal data fusion', *Neural Computation*, Vol. 32, No. 5, pp.829–864.

Gómez-Carmona, O., Sádaba, J. and Casado-Mansilla, D. (2022) 'Enhancing street-level interactions in smart cities through interactive and modular furniture', *Journal of Ambient Intelligence and Humanized Computing*, Vol. 13, No. 11, pp.5419–5432.

Hajzeri, A. (2021) 'The management of urban parks and its contribution to social interactions', *Arboricultural Journal*, Vol. 43, No. 3, pp.187–195.

Hangloo, S. and Arora, B. (2022) 'Combating multimodal fake news on social media: methods, datasets, and future perspective', *Multimedia Systems*, Vol. 28, No. 6, pp.2391–2422.

Hornecker, E., Hogan, T., Hinrichs, U. and Van Koningsbruggen, R. (2023) 'A design vocabulary for data physicalization', *ACM Transactions on Computer-Human Interaction*, Vol. 31, No. 1, pp.1–62.

Karasov, O., Külvik, M. and Burdun, I. (2021) 'Deconstructing landscape pattern: applications of remote sensing to physiognomic landscape mapping', *GeoJournal*, Vol. 86, No. 1, pp.529–555.

Lee, J-K. and Son, Y-H. (2022) 'Development of image classification model for urban park user activity using deep learning of social media photo posts', *Journal of the Korean Institute of Landscape Architecture*, Vol. 50, No. 6, pp.42–57.

Li, H., Goh, C., Hung, K. and Chen, J.L. (2018) 'Relative climate index and its effect on seasonal tourism demand', *Journal of Travel Research*, Vol. 57, No. 2, pp.178–192.

Liang, Z. and Li, X. (2023) 'What is a theme park? A synthesis and research framework', *Journal of Hospitality & Tourism Research*, Vol. 47, No. 8, pp.1343–1370.

Lin, Y-X. and Yang, Y. (2024) 'Understanding tourist restoration: an integrated framework from the perspective of environmental change', *Asia Pacific Journal of Tourism Research*, Vol. 29, No. 2, pp.144–160.

Lu, F., Chenwen, Y., Yuting, S. and Jing, Z. (2024) 'Generative design of plantscape based on generative adversarial network: a case study of the generation of flower border plan', *Landscape Architecture*, Vol. 31, No. 9, pp.59–68.

Ma, Y., Hao, Y., Chen, M., Chen, J., Lu, P. and Košir, A. (2019) 'Audio-visual emotion fusion (AVEF): a deep efficient weighted approach', *Information Fusion*, Vol. 46, pp.184–192.

Massaoudi, M.S., Abu-Rub, H. and Ghrayeb, A. (2023) 'Navigating the landscape of deep reinforcement learning for power system stability control: a review', *IEEE Access*, Vol. 11, pp.134298–134317.

Mfon, I. (2023) 'Aesthetic considerations in architectural design: exploring pleasure, arousal, and dominance', *International Journal of Research Publication and Reviews*, Vol. 4, No. 8, pp.923–935.

Rokhsaritalemi, S., Sadeghi-Niaraki, A. and Choi, S-M. (2023) 'Exploring emotion analysis using artificial intelligence, geospatial information systems, and extended reality for urban services', *IEEE Access*, Vol. 11, pp.92478–92495.

Serrano López, A.L., Freire Chaglla, S.A., Espinoza-Figueroa, F.E., Andrade Tenesaca, D.S. and Villafuerte Pucha, M.E. (2019) 'Modeling of tourist profiles with decision trees in a world heritage city: the case of Cuenca (Ecuador)', *Tourism Planning & Development*, Vol. 16, No. 5, pp.473–493.

Van Berkel, D.B., Tabrizian, P., Dorning, M.A., Smart, L., Newcomb, D., Mehaffey, M., Neale, A. and Meentemeyer, R.K. (2018) 'Quantifying the visual-sensory landscape qualities that contribute to cultural ecosystem services using social media and LiDAR', *Ecosystem Services*, Vol. 31, pp.326–335.

Wang, C., Zhang, J., Cao, J., Hu, H. and Yu, P. (2019) 'The influence of environmental background on tourists' environmentally responsible behaviour', *Journal of Environmental Management*, Vol. 231, pp.804–810.

Zheng, T., Yan, Y., Lu, H., Pan, Q., Zhu, J., Wang, C., Zhang, W., Rong, Y. and Zhan, Y. (2020) 'Visitors' perception based on five physical senses on ecosystem services of urban parks from the perspective of landsenses ecology', *International Journal of Sustainable Development & World Ecology*, Vol. 27, No. 3, pp.214–223.