Doctoral Dissertation

Integration of Multi-dimensional Assessment, Seasonal Plant-Specific Metrics, and 4D Modeling for Enhancing Urban Green Space Planning

HU ANQI

July 2024

Graduate School of Engineering,
Osaka University

博士学位論文

都市緑化計画を高度化するための
多次元評価、季節的な植物固有の指標および
4 次元モデリングの統合化

胡 安琪

2024 年 7 月

大阪大学
大学院工学研究科

Abstract

Urban green spaces are essential for promoting environmental, social, and economic sustainability in cities. However, existing methods for assessing urban green spaces often focus on single dimensions, failing to capture the complex nature of these spaces. The lack of objective and quantitative methods to address the multifunctionality and heterogeneity of urban green spaces, coupled with the seasonal variations and diverse plant species, leaves designers and planners relying on their prior knowledge for design and planning decisions.

This dissertation presents a comprehensive and innovative approach to assessing and modeling urban green spaces by integrating multi-source data, novel indicators, and advanced visualization technologies. The research introduces a comprehensive assessment framework that combines multiple data sources and evaluation metrics, such as the green view index (GVI) and green coverage ratio (GCR), to assess the spatial distribution, visibility, and composition of urban green spaces in a multi-dimensional manner. A case study of Osaka City demonstrates the framework's potential in revealing variations in green space provision and the influence of topography on green space distribution.

Furthermore, a multi-temporal urban green space vegetation visualization analysis framework is proposed, which integrates various data sources and advanced technologies to capture the temporal changes and seasonal variations in vegetation characteristics. Within this framework, the Seasonal Species-Specific Plant View Index (S3PVI) is introduced as a novel indicator to evaluate the visual importance and attractiveness of different plants in urban landscapes. The application of the multi-temporal analysis framework and the S3PVI indicator to the Sanshikisaido area in Suita, Osaka Prefecture, Japan, reveals the spatiotemporal differentiation patterns of the visibility characteristics of different plant types. A dataset containing 51 common plant species in urban environments is created to support the development and validation of the S3PVI indicator and the multi-temporal analysis framework.

The research also explores the application of advanced visualization technologies, such as neural radiance fields (NeRF) and Stable Diffusion, to generate dynamic and immersive 4D urban models that capture the temporal evolution and experiential qualities of urban green spaces. The integration of low-rank adaptation (LoRA) and ControlNet, built upon the data foundation laid in the multi-temporal analysis framework, enhances the visual quality and geometric consistency of the generated images.

The proposed frameworks and methodologies contribute to the development of more holistic, nuanced, and actionable approaches to understanding and managing urban green spaces. The research outcomes have the potential to support the creation of

sustainable and resilient urban environments by providing urban planners and policymakers with comprehensive insights and tools for evidence-based decision-making. However, further research and collaboration between diverse stakeholders are necessary to address the limitations, explore the full potential of the proposed frameworks, and develop an integrated system for urban green space planning and management.

Keywords: Urban green spaces; Multi-dimensional assessment; Seasonal Species-Specific Plant View Inde; Multi-temporal analysis; 4D modeling; Neural radiance fields; Stable Diffusion

Preface

This dissertation is the original work by Anqi Hu under the supervision of Prof. Nobuyoshi Yabuki. Two journal articles and two international conference proceedings related to this dissertation have been submitted or published. They are listed below.

Journal articles:

1. Hu, A., Yabuki, N., Fukuda, T., Kaga, H., Takeda, S., & Matsuo, K. (2023). Harnessing Multiple Data Sources and Emerging Technologies for Comprehensive Urban Green Space Evaluation. *Cities,* 143, 104562. https://doi.org/10.1016/j.cities.2023.104562

2. Hu, A., Yabuki, N., & Fukuda, T. A Multi-temporal Framework for Urban Green Space Vegetation Visualization and Analysis Using Deep Learning and 3D Reconstruction. *Cities*, Under review.

International conference proceedings:

1. Hu, A., Yabuki, N., & Fukuda, T. (2023). Development of a Method for Assessing the View Index of Plants of Interest Using Deep Learning, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 585-594. https://doi.org/10.52842/conf.caadria.2023.1.585

2. Hu, A., Yabuki, N., & Fukuda, T. (2024). Generating 4D Plant Models for Virtual Reality Environments Using the Instant Neural Graphics Primitives and Stable Diffusion Model. Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 3,411-420.

Acknowledgment

This doctoral dissertation would not have been possible without the guidance, support, and assistance of numerous individuals to whom I am deeply indebted.

First and foremost, I extend my profound gratitude to my Ph.D. supervisor, Professor Nobuyoshi Yabuki, for his unwavering encouragement and guidance. His consistent and illuminating instruction has been instrumental in shaping this research into its present form. Professor Yabuki not only provided continuous support but also granted me the freedom to pursue various research projects. I am particularly grateful for the opportunity he afforded me to participate in international conferences, which proved to be an invaluable experience during my doctoral studies.

I am equally indebted to my co-supervisor, Associate Professor Tomohiro Fukuda, for his patient instruction, insightful criticism, and expert guidance. Dr. Fukuda's astute comments on all my papers have been immensely beneficial to my academic growth.

A special debt of gratitude is owed to the dissertation committee members for dedicating their time to reading this dissertation and providing valuable feedback and recommendations. I am particularly thankful to Professor Masanobu Kii for serving as the vice referee for my Ph.D. defense.

My heartfelt thanks go to my parents, Wang Liyun and Hu Yaoping, and my family for their love, trust, and unconditional support over the years. I am also sincerely grateful to my friends Zhang Jiaxin, Zhao Jiapei, Chen Sihua, and Zhou Yuxin, who offered their time, lent an ear to my concerns, and helped me navigate through challenging periods of my research. I extend my appreciation to my senior colleagues, Li Yunqin and Zhang Jiaxin, who have since graduated, for their constructive suggestions.

I would like to acknowledge the support of my laboratory colleagues: Yu, Mugita, Yang, Nishimoto, Ogura, Tsurunaga, Futamura, Yoshimura, Shimizu, Soe, Liang, Shi, Shirahase, Fukuda, Hashizume, and others, for their assistance and camaraderie in daily life.

My sincere thanks also go to Ms. Sun, Ms. Wang, Professor Kaga, Professor Takeda, and Dr. Matsuo, without whose support I would not have had the opportunity to study abroad and pursue my doctoral degree.

I would also like to express my appreciation to my former best friend Zhang, for the support and companionship that contributed to my personal growth.

Lastly, I am grateful to myself for persevering through challenges and never giving up.

The memories of these three years will remain etched in my mind as the most precious experiences of my life.

Table of Contents

Abstract .... **i**

Preface .... **iii**

Acknowledgment .... **v**

List of Figures .... **xi**

List of Tables .... **xiii**

List of Abbreviations .... **xv**

Chapter 1 Introduction **1**

1.1 Research background and problem statements 1

1.2 Research objectives 2

1.2.1 Main objective 2

1.2.2 Specific goals 3

1.2.3 Practical applications 4

1.2.4 Theoretical contributions 4

1.3 Research scopes 5

1.3.1 Data focus 5

1.3.2 Data sources 5

1.3.3 Methodology 5

1.4 Overview of the dissertation 6

Chapter 2 Literature review **9**

2.1 Limitations and potential of visual characteristics assessment methods for urban green spaces 9

2.1.1 Limitations of traditional assessment methods 9

2.1.2 Potential of computer vision and semantic segmentation techniques 10

2.2 Relationship between vegetation visual characteristics and human perception 10

2.2.1 Importance of visual diversity and individual plant species aesthetics 10

2.2.2 Incorporating aesthetic considerations in urban green space assessment 10

2.3 Application and limitations of street view images in urban vegetation analysis 11

2.3.1 Current applications and challenges 11

2.3.2 Advances in data preprocessing and integration 11

2.4 Spatiotemporal dynamic analysis of visual characteristics of urban green spaces 12

2.5 Innovations in urban green space assessment and modeling 12

2.6 Summary of research gaps and goals 13

Chapter 3 Comprehensive assessment framework for urban green spaces 17

3.1 Overview 17

3.2 Method and materials 18
3.2.1 Research framework 18
3.2.2 GCR calculation 20
3.2.3 GVI calculation 21
3.2.4 Green space type classification 23

3.3 Study area and data sources 23
3.3.1 Study area 23
3.3.2 Data sources 25

3.4 Results 26
3.4.1 Data integration results 26
3.4.2 Image classification results 30
3.4.3 Correlation between topography and green space 31
3.4.4 Data on green space in four groups of elementary schools 32

3.5 Discussion 33
3.5.1 Distribution and characteristics of green space types 33
3.5.2 Topographical factors influencing urban greening 34
3.5.3 Green space distribution among elementary school groups 34
3.5.4 Limitations and future research directions 35

3.6 Summary of this chapter 36

Chapter 4 A multi-temporal evaluation framework for the S3PVI 39

4.1 Overview 39

4.2 Proposed framework 40
4.2.1 Data collection and preprocessing 41
4.2.2 Standardized view generation 43
4.2.3 Plant visual feature evaluation based on S3PVI 45
4.2.4 S3PVI calculation for plant visual feature quantification 50
4.2.5 Seasonal change analysis of plant visual features .... **Error! Bookmark not defined.**

4.3 Experiments and results 50
4.3.1 Visual feature analysis of real street vegetation 50
4.3.2 Multi-temporal visualization evaluation of virtual park vegetation design schemes 62

4.4 Discussion 76
4.4.1 Advantages and limitations of the multi-temporal urban green space vegetation visualization analysis framework 76
4.4.2 The significance of the S3PVI for quantitative evaluation 77

4.4.3 Application of the multi-temporal analysis framework in empirical case studies 77
4.4.4 Multi-temporal vegetation visualization analysis supporting 4D vegetation landscape modeling 78
4.4.5 Improvement directions and application extensions 79
4.5 Summary of this chapter 79
Chapter 5 Plant landscape modeling: integrating dynamics and techniques 81
5.1 Overview 81
5.2 Proposed framework 82
5.2.1 Technical framework construction 82
5.2.2 Model generation pipeline 84
5.2.3 The main process of the system 85
5.2.4 Model evaluation metrics 87
5.3 Experiments and results 88
5.3.1 Experimental setup 88
5.3.2 Data and target plant selection 89
5.3.3 Seasonal variation results 90
5.3.4 ControlNet model evaluation 92
5.3.5 Image quality evaluation 93
5.3.6 Seasonal variation results for streetscapes 94
5.4 Discussion 95
5.4.1 Limitations of controlled conditions in digital tree modeling 95
5.4.2 Extending the framework to multi-plant scenes and ecosystems 97
5.4.3 Reducing dependence on extensive real image data 97
5.4.4 Converting image-space representations into explicit 3D models 97
5.4.5 Integration with urban simulation tools and decision support systems 98
5.4.6 Complementary relationship between 4D modeling and S3PVI 98
5.4.7 Future work and challenges 98
5.5 Summary of this chapter 99
Chapter 6 Conclusion 101
6.1 Summary 101
6.1.1 Comprehensive assessment framework for urban green spaces 101
6.1.2 Multi-temporal urban green space vegetation visualization analysis framework and S3PVI 101
6.1.3 Application of advanced visualization technologies 102
6.2 Research contributions 103
6.2.1 Comprehensive assessment framework 103
6.2.2 Multi-temporal urban green space vegetation visualization analysis framework and S3PVI 103

6.2.3 Integration of advanced visualization technologies 103

6.3 Limitations and future work 104

References 109

List of Figures

Figure 1.1 Overview of the main concepts in each chapter. 6
Figure 3.1 Workflow of the urban green space assessment framework. 19
Figure 3.2 Schematic diagram of the PSPNet image segmentation model. 21
Figure 3.3 GSV images captured at a sample site in the study area from six different directions (0°, 60°, 120°, 180°, 240°, and 300°). 22
Figure 3.4 Geographical location of the study area. 24
Figure 3.5 Distribution of NDVI in Osaka city. 26
Figure 3.6 Spatial distribution of average GVI in Osaka. 27
Figure 3.7 Comparison of GVI and GCR in Osaka's elementary school districts. 28
Figure 3.8 Relationship between GVI and GCR in four groups of elementary schools in Osaka. 29
Figure 3.9 Classification of green space types in Osaka based on GVI. 30
Figure 3.10 Classification of elementary school districts based on terrain distribution. 31
Figure 3.11 Average GVI values for different green space types in Osaka. 32
Figure 3.12 Proportion of different green space types in Osaka's elementary school districts based on GSV image analysis. 33
Figure 4.1 Workflow diagram of the multi-temporal urban green space vegetation visualization analysis framework. 41
Figure 4.2 Process of 3D scene reconstruction using SfM and 3D Gaussian splatting 42
Figure 4.3 Standardized imaging method for plant analysis. 44
Figure 4.4 Comparison of street view images before and after preprocessing. 45
Figure 4.5 Examples of original and segmented images. 49
Figure 4.6 Vegetation distribution along Sanshikisaido, Suita City. 51
Figure 4.7 Detailed vegetation map and camera placement in zone 1 of Sanshikisaido. 52
Figure 4.8 Detailed vegetation map and camera placement in zone 2 of Sanshikisaido. 53
Figure 4.9 Detailed vegetation map and camera placement in zone 3 of Sanshikisaido. 54
Figure 4.10 Detailed vegetation map and camera placement in zone 4 of Sanshikisaido. 55
Figure 4.11 Temporal analysis of vegetation presence at location 1 (2010-2022) 57
Figure 4.12 Temporal analysis of vegetation presence at location 2 (2010-2022) 58
Figure 4.13 Temporal analysis of vegetation presence at location 25 (2010-2022) 59
Figure 4.14 Temporal analysis of vegetation presence at location 45 (2010-2022) 60
Figure 4.15 Seasonal change curves of S3PVI values for major plant species 61
Figure 4.16 Virtual park design layout for multi-temporal vegetation visualization. 63
Figure 4.17 Location map of the virtual camera setup for the virtual park. 64
Figure 4.18 Seven planting design schemes for the virtual park 65
Figure 4.19 Seasonal transformation of vegetation in Scheme 1. 67
Figure 4.20 Seasonal transformation of vegetation in Scheme 2. 68
Figure 4.21 Seasonal transformation of vegetation in Scheme 3. 69
Figure 4.22 Seasonal transformation of vegetation in Scheme 4. 70

Figure 4.23 Seasonal transformation of vegetation in Scheme 5. 71
Figure 4.24 Seasonal transformation of vegetation in Scheme 6. 72
Figure 4.25 Seasonal transformation of vegetation in Scheme 7. 73
Figure 4.26 Seasonal S3PVI values for seven planting design schemes in the virtual park 75
Figure 5.1 4D generation system for plant landscape modeling. 86
Figure 5.2 Workflow for training LoRA model using multi-temporal vegetation visualization. ...90
Figure 5.3 The state of the target plant in the VR environment in all seasons. 91
Figure 5.4 Comparative results of ControlNet models for edge detection, depth estimation, and semantic segmentation. 93
Figure 5.5 Comparative analysis of seasonal transitions in VR environment. 95
Figure 5.6 Example of image quality degradation in complex real-world scenes....................96

List of Tables

Table 3.1 Data sources used in the study..25
Table 4.1 Virtual camera parameter settings for standardized view generation 43
Table 4.2 IoU scores for common urban plant species in semantic segmentation..........................45
Table 5.1 Performance comparison between NeRF and 3D Gaussian splatting.............................83
Table 5.2 Equipment, software, and parameters used..88
Table 5.3 Parameters used in Stable Diffusion for seasonal variations. 92
Table 5.4 Comparative analysis of image quality metrics. 94

List of Abbreviations

Acronyms	Definitions
CNN	Convolutional neural network
GCR	Green coverage rate
GIS	Geographic information system
GUI	Graphical user interface
GVI	Green view index
Instant-ngp	Instant neural graphics primitives
IoT	Internet of things
LiDAR	Light detection and ranging
LoRA	Low-rank adaptation
LPIPS	Learned perceptual image patch similarity
mIoU	Mean intersection over union
MR	Mixed reality
MSE	Mean square error
NeRF	Neural radiance fields
NDVI	Normalized difference vegetation index
POI	Point of interest
PPM	Pyramid pooling module
PSNR	Peak signal-to-noise ratio
S3PVI	Seasonal species-specific plant view index
SfM	Structure from motion
SSIM	Structural similarity index
VR	Virtual reality

Chapter 1
Introduction

1.1 Research background and problem statements

Urban green spaces, including parks, gardens, street greenery, and urban forests, play a vital role in promoting the environmental, social, and economic sustainability of cities (Wolch et al., 2014). They provide numerous ecosystem services, such as improving air quality, regulating urban temperatures, managing stormwater runoff, and supporting biodiversity (Elmqvist et al., 2015). Moreover, urban green spaces contribute to the physical and mental well-being of city dwellers by offering opportunities for recreation, social interaction, and stress relief (Hartig et al., 2014).

As cities continue to expand and densify, the importance of urban green spaces in creating livable and resilient urban environments has become increasingly recognized by researchers, policymakers, and the public alike (Lovell & Taylor, 2013). However, the rapid urbanization process also poses challenges for the planning, design, and management of urban green spaces (Haaland & Van Den Bosch, 2015). Urban green spaces are often fragmented, unevenly distributed, and subject to competing land-use demands (Pauleit et al., 2019).

To ensure that urban green spaces can effectively deliver their intended benefits, it is crucial to accurately assess and monitor their quantity, quality, and distribution (Xu et al., 2020). Traditional assessment methods, such as the green view index (GVI) (Aoki et al., 1985), green coverage ratio (GCR) and the normalized difference vegetation index (NDVI), have been widely used to quantify the extent and density of urban vegetation (Gupta et al., 2012; A. Hu, Yabuki, Fukuda, et al., 2023; Weier & Herring, 2000; Zhu et al., 2023). However, these methods often fail to capture the multidimensional characteristics of urban green spaces, such as their ecological functions, aesthetic qualities, and social benefits (Cilliers & Cilliers, 2015; Hoyle et al., 2017).

Furthermore, urban green spaces are dynamic systems that undergo seasonal changes and long-term evolution (Han et al., 2023; Wellmann et al., 2020). The visual appearance and ecological functions of urban vegetation vary significantly throughout the year, influenced by factors such as plant phenology, climate conditions, and management practices. However, most existing assessment methods rely on static or single-temporal data, failing to account for the temporal dynamics of urban green spaces (Zhou et al., 2016).

Another challenge in urban green space assessment is the limited integration of quantitative measures with aesthetic considerations (R. Wang et al., 2016). The visual quality and aesthetic appeal of urban green spaces play a crucial role in shaping public perception, preference, and well-being. Aesthetically pleasing urban green spaces not only enhance the emotional well-being of urban residents but also contribute to the economic value of the surrounding areas (Xiao et al., 2017). However, current assessment methods often focus on the overall greenness or vegetation health, neglecting the aesthetic qualities and visual diversity of individual plant species (A. Hu, Yabuki, & Fukuda, 2023).

Currently, urban planners and designers often rely on subjective knowledge and experience when selecting and designing urban green spaces, which may lead to suboptimal outcomes (Riechers et al., 2019). There is a need for objective and quantitative assessment methods and indicators that can guide evidence-based decision-making in urban green space planning and design (Noland et al., 2017).

To address these challenges, there is a pressing need for innovative approaches and methodologies that can provide a more comprehensive, nuanced, and actionable understanding of urban green spaces. This research aims to fill this gap by developing a multi-dimensional assessment framework, introducing novel indicators, and exploring advanced visualization technologies for urban green space assessment and modeling.

1.2 Research objectives

1.2.1 Main objective

The main objective of this research is to develop a comprehensive and innovative approach for assessing and modeling urban green spaces, integrating multi-source data, novel indicators, and advanced visualization technologies to support evidence-based decision-making in urban planning and management.

This approach aims to provide a multi-dimensional understanding of urban green spaces, with a primary focus on their visual and spatio-temporal characteristics. It encompasses:

The integration of spatial distribution, visibility, and composition metrics to offer a more holistic assessment of urban green spaces.

The introduction of novel indicators that capture the seasonal dynamics and visual importance of different plant species in urban landscapes.

The exploration of advanced visualization technologies to generate dynamic and immersive representations of urban vegetation over time.

By combining these elements, this study strives to create a relatively comprehensive framework for analyzing and visualizing the visual aspects and temporal evolution of urban green spaces. This focus on visual and temporal dimensions is intended to complement existing research on other crucial aspects of urban green spaces, such as biodiversity, ecosystem services, and social functions. While these other aspects are undoubtedly important, they are beyond the scope of this particular study.

The proposed approach aims to bridge the gap between quantitative assessment and qualitative experience of urban green spaces, providing urban planners and designers with tools to better understand, communicate, and enhance the visual and temporal qualities of urban vegetation. Through this focused yet comprehensive approach, this research contributes to the broader field of urban green space studies while maintaining a specific emphasis on visual and spatio-temporal aspects.

1.2.2 Specific goals

1) Developing a comprehensive assessment framework: The integration of multi-source data and evaluation metrics enables the multi-dimensional assessment of spatial distribution, visibility, and composition of urban green spaces.
2) A multi-temporal visualization analysis framework: Capturing temporal changes and seasonal variations in vegetation characteristics, this framework aids in the analysis of urban green spaces over different time scales.
3) Advanced visualization technology in dynamic and immersive 4D urban models: Application of these technologies captures the temporal evolution and experiential qualities of urban green spaces, providing a comprehensive understanding.
4) The interaction between topography and green spaces: Studying this interaction to better understand the complex relationship between green spaces and the urban environment.
5) The potential of street-view data: Intelligent classification techniques unlock the potential of street-view data in evaluating and understanding the multifunctionality and distribution patterns of urban green spaces.

1.2.3 Practical applications

The outcomes of this research may have valuable practical implications for urban planning, design, and management. The proposed assessment framework and novel indicators can provide urban planners and policymakers with a more comprehensive and nuanced understanding of the spatial distribution, composition, and quality of urban green spaces. This knowledge can inform evidence-based decision-making processes, such as identifying areas of green space deficiency, prioritizing green space interventions, and optimizing the allocation of resources for green space development and maintenance.

Moreover, the multi-temporal vegetation visualization analysis framework and the seasonal species-specific plant view index (S3PVI) indicator can guide plant selection and configuration decisions in urban landscape design. By considering the seasonal changes and visual characteristics of individual plant species, designers can create visually appealing and engaging urban green spaces that provide year-round aesthetic benefits and enhance the well-being of urban residents. The quantitative analysis results provided by this research can serve as objective references for urban planners and designers, complementing their subjective knowledge and experience in green space planning and design.

The application of advanced visualization technologies, such as neural radiance fields (NeRF) and Stable Diffusion, in generating dynamic and immersive 4D urban models can revolutionize the way urban green spaces are planned, designed, and communicated. These technologies enable stakeholders, including planners, designers, policymakers, and the public, to visualize and experience the temporal evolution and experiential qualities of urban green spaces. This can facilitate more effective collaboration, public participation, and consensus-building in urban planning processes.

1.2.4 Theoretical contributions

This research aims to contribute to the fields of urban ecology, landscape assessment, and computational modeling. The development of the comprehensive assessment framework and the S3PVI indicator can enhance the understanding of the multidimensional characteristics and temporal dynamics of urban green spaces. By integrating multi-source data and considering both quantitative and qualitative aspects of urban vegetation, this research can provide a more holistic and nuanced perspective on the assessment of urban green spaces.

The exploration of advanced visualization technologies and 4D modeling techniques can expand the boundaries of traditional urban landscape representation and analysis. The integration of NeRF and Stable Diffusion in generating dynamic and immersive urban models can open up new possibilities for studying the temporal evolution and experiential qualities of urban green spaces. This research can contribute to the development of innovative computational methods and tools for urban landscape assessment and modeling.

Furthermore, the investigation of the interplay between topography and green spaces can deepen the understanding of the complex relationship between urban green spaces and the built environment. By considering the influence of topographical factors on the distribution, composition, and functionality of urban green spaces, this research may offer valuable insights into the ecological and social dimensions of urban landscapes.

1.3 Research scopes

1.3.1 Data focus

This research focuses on assessing and modeling urban green spaces, including parks, gardens, street greenery, and urban forests. The primary data sources used in this research are multi-source geospatial data, including remote sensing imagery and street-view images.

1.3.2 Data sources

The data used in this research are obtained from various sources, including:

1) Remote sensing imagery: High-resolution satellite imagery (e.g., Sentinel-2) and aerial photographs are used to assess the spatial distribution and composition of urban green spaces.
2) Street-view images: Large-scale street-view image datasets, such as Google Street View, are used to evaluate the visual characteristics and seasonal variations of urban vegetation from a pedestrian perspective.
3) Geospatial data: Various geospatial datasets, such as land use/land cover maps, administrative boundaries, and road networks, are used to provide contextual information and support the analysis of urban green spaces.

1.3.3 Methodology

This research employs a range of advanced methodologies and techniques, including:

1) Geospatial analysis: Geospatial analysis techniques are used to assess the spatial distribution, composition, and connectivity of urban green spaces.
2) Computer vision: Computer vision techniques, such as semantic segmentation and image classification, are used to extract and analyze vegetation information from street-view images and remote sensing imagery.
3) Deep learning: Deep learning models, such as Stable diffusion and NeRF, are used to generate realistic and dynamic 4D urban models.
4) 3D reconstruction and visualization: 3D reconstruction techniques (Mouragnon et al., 2006), such as structure from motion (SfM) are used to generate 3D point clouds and models of urban green spaces.

1.4 Overview of the dissertation

This dissertation is separated into 6 chapters, with the main concepts of each chapter illustrated in Figure 1.1.

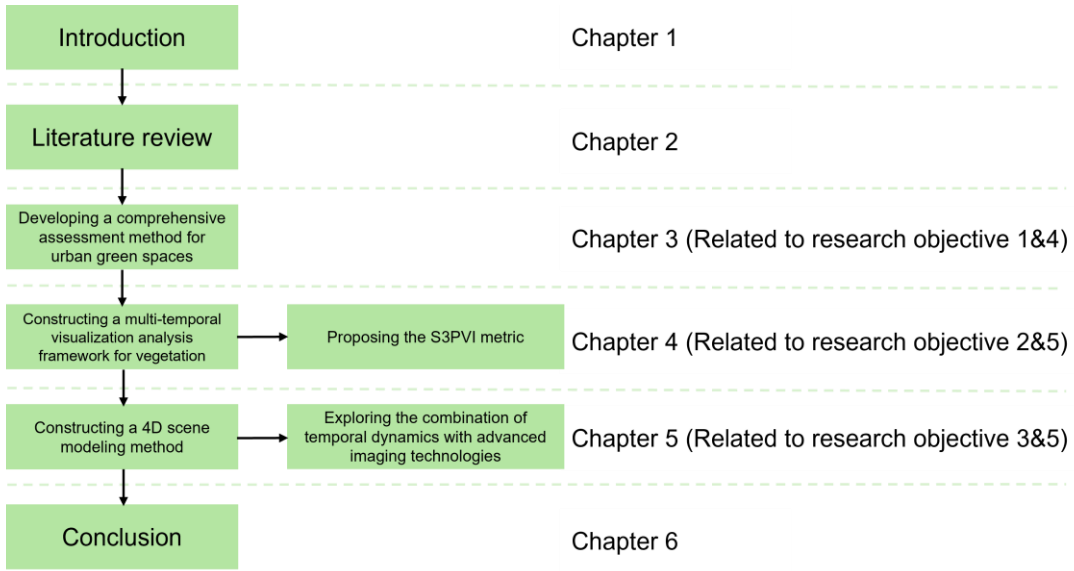


Figure 1.1 Overview of the main concepts in each chapter.

The next chapter reviews the literature related to this dissertation. Subsequently, Chapter 3 develops a comprehensive assessment framework that integrates multi-source data and evaluation metrics to assess the composition of urban green spaces in a multi-dimensional manner. This chapter addresses the first and fourth research objectives. In Chapter 4, the assessment metrics from Chapter 3 are refined to better describe the visual characteristics and attractiveness of different plants in urban landscapes. Additionally, a multi-temporal visualization analysis framework for urban green spaces is constructed, relating to the second and fifth research objectives. Chapter 5 explores the application of advanced visualization technologies to generate dynamic and immersive 4D urban models, capturing the temporal evolution and experiential quality of urban green spaces. This chapter is associated with the third and fifth research objectives. The final chapter, Chapter 6, summarizes the findings from each chapter and provides recommendations for future research and practical applications. The detail of each chapter is further explained:

Chapter 1: Introduction

Chapter 1 provides an overview of the research background, problem statements, objectives, significance, scopes, and framework. It sets the context for the research and highlights the key challenges and opportunities in urban green space assessment and modeling.

Chapter 2: Literature review

Chapter 2 presents a comprehensive review of the existing literature on urban green space assessment, landscape visualization, and computational modeling. It identifies the research gaps and limitations in current approaches and highlights the need for innovative methodologies and indicators.

Chapter 3: Comprehensive assessment framework for urban green spaces

Chapter 3 introduces the comprehensive assessment framework for urban green spaces, which integrates multi-source data and evaluation metrics (GVI, GCR) to assess the spatial distribution, visibility, and composition of urban green spaces in a multi-dimensional manner. It demonstrates the application of the framework using a case study and discusses the implications for urban planning and management.

Chapter 4: A multi-temporal evaluation framework for the S3PVI

Chapter 4 focuses on the development and application of the S3PVI as a novel indicator to evaluate the visual importance and attractiveness of different plants in urban landscapes. It also presents the multi-temporal urban green space vegetation visualization analysis framework, which captures the temporal changes and seasonal variations in vegetation characteristics.

Chapter 5: Plant landscape modeling: integrating dynamics and techniques

Chapter 5 explores the application of advanced visualization technologies, in generating dynamic and immersive 4D urban models to capture the temporal evolution and experiential qualities of urban green spaces. It demonstrates the potential of these technologies for enhancing the understanding and communication of urban green space dynamics.

Chapter 6: Conclusion

Chapter 6 summarizes the main findings and contributions of the research, discusses the implications for urban planning and management, and provides recommendations for future research and practical applications. It also highlights the limitations of the research and identifies potential avenues for further investigation.

In summary, this dissertation presents a comprehensive and innovative approach to urban green space assessment and modeling, integrating multi-source data, novel indicators, and advanced visualization technologies. By addressing the research gaps and challenges identified in the literature, this research aims to contribute to the development of more sustainable, resilient, and livable urban environments.

Chapter 2
Literature review

2.1 Limitations and potential of visual characteristics assessment methods for urban green spaces

2.1.1 Limitations of traditional assessment methods

A significant limitation of current urban green space assessment methods is their inability to capture the temporal changes and seasonal variations in vegetation characteristics (Dutta et al., 2022; Shiraishi & Terada, 2024). Urban vegetation exhibits distinct phenological patterns, with changes in foliage color, density, and overall appearance throughout the year (Kuper, 2015). These temporal dynamics play a crucial role in shaping the visual quality and aesthetic appeal of urban green spaces (Hoyle et al., 2017). However, most existing assessment techniques rely on static or single-temporal data, failing to account for the dynamic nature of urban vegetation (Zhou et al., 2016).

Moreover, current urban green space planning and design practices often rely on subjective experiences and knowledge of planners and designers when selecting and arranging plant species (Beer et al., 2003). While expert knowledge is valuable, there is a lack of quantitative tools and evidence-based approaches to guide plant selection and configuration decisions (Daniels et al., 2018). This gap highlights the need for more objective and data-driven methods to assess the visual characteristics of individual plant species and their contribution to the overall aesthetic quality of urban green spaces.

2.1.2 Potential of computer vision and semantic segmentation techniques

With the development of computer vision technology, scholars have explored the use of big data to assess the visual characteristics of urban green spaces. Seiferling et al. (2017) and Hu et al. (2023) utilized semantic segmentation and deep learning techniques to analyze street greening levels and the spatial distribution of trees. Semantic segmentation is a key technique in computer vision that has been successfully applied to urban scene understanding and vegetation mapping (Sodjinou et al., 2022). By assigning pixel-level labels to different object categories, semantic segmentation enables fine-grained analysis of urban landscapes, including the delineation of individual plant species.

Object-based classification techniques have shown promise in detecting, delineating, and classifying urban plant species from very high-resolution satellite imagery (Sicard et al., 2023). By selecting relevant spectral and texture-based features for each plant species, these methods can achieve high classification accuracies. However, challenges remain due to the complexity of the urban environment, the diversity of species, and the spatial proximity between trees.

2.2 Relationship between vegetation visual characteristics and human perception

2.2.1 Importance of visual diversity and individual plant species aesthetics

Understanding the relationship between vegetation visual characteristics and human perception is crucial for guiding the assessment and design of urban green spaces. Recent studies have highlighted the importance of visual attributes in shaping people's appreciation and restorative experiences in urban green spaces (Du et al., 2016). Visual diversity, colorful vegetation, and a mix of evergreen and deciduous plants have been found to enhance aesthetic preferences and perceived restorative potential.

The aesthetic appeal of individual plant species is an essential factor in creating visually pleasing and engaging urban environments (Carrus et al., 2015; Ulrich, 1986). Different plant species possess unique visual characteristics, such as foliage color, texture, and growth habits, which contribute to the overall aesthetic diversity of urban green spaces (Lindemann-Matthies & Brieger, 2016). However, most current assessment methods focus on the quantitative aspects of urban vegetation (Stessens et al., 2020), such as coverage ratio or vegetation health, while neglecting the aesthetic qualities and visual diversity of individual plant species.

2.2.2 Incorporating aesthetic considerations in urban green space assessment

Incorporating the aesthetic considerations of individual plant species into urban green space assessment is crucial for several reasons. First, visually attractive and diverse plant palettes can enhance the economic value of urban areas by increasing property values and attracting businesses and visitors (X. Wang et al., 2016). Second, engaging with aesthetically pleasing vegetation has been shown to provide psychological benefits, such as stress reduction and improved well-being (Bielinis et al., 2018). Finally, understanding the aesthetic preferences of the public can inform plant selection and configuration decisions, ensuring that urban green spaces are designed to meet the needs and expectations of the community.

Studies have proposed combining visual diversity assessments with quantitative indicators to assess perceived aesthetic quality. Kuper (2015) used a visual rating approach to score the plant aesthetics of parks in terms of color, diversity, rhythm, texture, and other aspects. By integrating qualitative and quantitative measures, these studies underscore the importance of considering key vegetation characteristics when assessing and designing visually appealing urban green spaces. The development of vegetation visual characteristic indices that capture the aesthetic diversity and seasonal dynamics of individual plant species is an emerging area of research that could contribute to more effective urban planning and management.

2.3 Application and limitations of street view images in urban vegetation analysis

Street view images provide a valuable data source for urban green space vegetation analysis by capturing urban environments from a pedestrian perspective. Researchers have utilized street view images to analyze the spatial distribution characteristics of urban green spaces (Kameoka et al., 2022; Li, 2021). The use of point of interest (POI) data and semantic segmentation models has further enabled the classification and analysis of urban green spaces based on street-view data (Chen et al., 2018; X. Zhang et al., 2017).

2.3.1 Current applications and challenges

However, there are limitations in the current application of street view images in urban vegetation analysis. Most studies use single-temporal images, making it difficult to reflect seasonal changes in vegetation (Anguelov et al., 2010). Moreover, due to the limitations of vehicle-mounted equipment, shooting angles and positions are not easily adjustable. Differences in image quality and resolution from different sources also pose challenges for image analysis.

2.3.2 Advances in data preprocessing and integration

To address these limitations, Xia et al. (2021) developed a method based on semantic segmentation processing of street view images to calculate the green view index of urban streets and proposed the panoramic view green view index for measuring visible street-level greenery. 3D point cloud data integrated with 3D Gaussian splatting significantly enhances the preprocessing of visual data for urban environmental assessments. Initially developed for computer graphics, this technique now improves the quality and consistency of urban street view images, which is crucial for accurate greenery assessments. Recent advances, such as those by Oh et al. (2024), demonstrate its ability to combine light detection and ranging (LiDAR) data with visual cues to produce highly accurate urban scene renderings. This method not only ensures consistent image properties but also aids urban planning and landscape architecture by providing detailed, species-specific insights into urban greenery.

2.4 Spatiotemporal dynamic analysis of visual characteristics of urban green spaces

The visual characteristics of urban green spaces exhibit significant spatiotemporal heterogeneity. Multi-temporal remote sensing images have been used to analyze the dynamic changes in urban green spaces (Dutta et al., 2022; Shiraishi & Terada, 2024). However, limited by the spatial resolution of remote sensing data, these studies have difficulty in refining the internal vegetation conditions of urban green spaces and their visual characteristics from a pedestrian perspective.

Recent studies have begun to explore the use of multi-temporal street view images to analyze the seasonal dynamics of urban green spaces. Song et al. (2018) proposed a dynamic method to assess urban greenspace exposure with the integration of mobile-phone locating-request data and high-spatial-resolution remote sensing images. Liang et al. (2023) introduced an embedding-driven clustering approach that integrates both physical and perceptual attributes to infer the spatial structure of the visual environment and investigate its spatiotemporal evolution.

However, there is currently a lack of systematic analysis of the multi-temporal and multi-perspective dynamics of urban green spaces, and the refined characterization of vegetation types needs to be strengthened. Integrating remote sensing data with social sensing data has shown promise in determining urban sprawl and its impact on sustainable urban development (Shao et al., 2021).

2.5 Innovations in urban green space assessment and modeling

To address the limitations of existing methods, innovative approaches have been proposed for urban green space assessment and dynamic urban modeling. Comprehensive assessment frameworks that integrate multiple data sources, including remote sensing imagery, street-level photographs, and geospatial data, have been developed to provide a holistic evaluation of urban green spaces (Pulighe et al., 2016).

Advanced visualization technologies have shown potential for generating realistic and immersive urban models that capture the temporal dynamics and experiential qualities of urban green spaces (Lu et al., 2023; Yao et al., 2024). These technologies enable the creation of photorealistic and temporally aware visualizations that support more effective communication and collaboration in urban planning and management processes.

The integration of urban sensing technologies, such as internet of things (IoT) devices and mobile crowdsourcing, has enabled the collection of real-time and fine-grained data on urban dynamics (Kandt & Batty, 2021). These data sources can be assimilated into urban models using data fusion and machine learning techniques, enhancing the accuracy and responsiveness of urban models.

Despite the progress in urban green space assessment, current methods still have notable limitations. Most approaches focus on the quantitative aspects of urban greenery, such as coverage ratio or vegetation health, while neglecting the aesthetic qualities and visual diversity of different plant species (Lindemann-Matthies & Brieger, 2016). The lack of refined characterization of vegetation species diversity has not been fully addressed. Some scholars have attempted to identify plant species from street view images using computer vision techniques (Kotowska et al., 2021) and measure community-level species richness by combining remote sensing data (Chavan, 2023), but accurate recognition in complex urban environments remains challenging. Future research should further explore the fusion of multi-source heterogeneous data and integrate knowledge-driven modeling approaches to better represent the visual diversity of urban green spaces.

2.6 Summary of research gaps and goals

The existing literature highlights several research gaps and limitations in current urban green space assessment methods. Traditional evaluation measures, such as the proportion or volume of green space, tend to oversimplify the evaluation process, neglecting the composite nature of urban green spaces (Balram & Dragievi, 2005; Grunewald & Bastian, 2015). Conventional assessments that rely solely on single metrics fail to capture the full range of functions and services that green spaces offer. For example, the GVI survey, which primarily relies on street view data (J. Zhang & Hu, 2022) , disregards green spaces that extend beyond streets, limiting the evaluation to a one-dimensional perspective.

Researchers emphasize the need to transcend the evaluation of isolated indicators and explore their interplay and mutual influences (Carmen et al., 2020; James et al., 2009). Capturing, manipulating, and measuring urban green spaces' functions and services present pressing challenges that require innovative approaches and data integration across disciplines (Luederitz et al., 2015). Moreover, the complex relationship between green spaces and the urban environment, influenced by topographical characteristics, requires further exploration.

One of the most significant research gaps identified in the literature is the lack of methods to capture and characterize the temporal changes and seasonal variations in vegetation characteristics. Most existing assessment techniques rely on static or single-temporal data, failing to account for the dynamic nature of urban vegetation. This limitation hinders the ability to understand and represent the aesthetic qualities and visual diversity of urban green spaces throughout the year.

Another research gap is the limited integration of quantitative measures with aesthetic considerations in assessing urban green spaces. Current methods often focus on the overall greenness or vegetation health, neglecting the aesthetic qualities and visual diversity of individual plant. Incorporating the aesthetic aspects of individual plants is crucial for creating visually appealing and engaging urban environments that provide psychological benefits and enhance the well-being of urban residents.

Furthermore, there is a lack of advanced visualization technologies and immersive 4D modeling techniques that can effectively capture the temporal evolution and experiential qualities of urban green spaces. Developing realistic and dynamic 4D models of urban vegetation is essential for understanding the multi-temporal and multi-perspective dynamics of urban green spaces and facilitating effective communication and collaboration in urban planning and management processes.

Through a systematic review of the limitations in current urban green space assessment methods, particularly in terms of survey content, data timeliness, and aesthetic perspectives, this study identifies the research gaps that warrant further exploration, which also form the main research objectives of this dissertation (as outlined in Chapter 1):

1) A comprehensive assessment framework for urban green spaces that integrates multi-source data to enable a multi-dimensional characterization and quantitative analysis of the spatial distribution, visibility, and composition of urban green spaces.
2) A multi-temporal visualization analysis method for urban green space vegetation, introducing seasonal variation characteristics and visual aesthetic indicators, and a new vegetation visualization index to reveal the spatiotemporal differentiation patterns of urban green space landscapes.
3) The application of advanced visualization technologies in urban green space modeling to generate dynamic and immersive urban green space landscapes, expanding the expression capabilities of spatiotemporal characteristics and human-centered experiences of urban green spaces.
4) The influence of natural geographical elements, such as topography, on the spatial pattern of urban green spaces, and a systematic characterization of the formation mechanisms and differentiation characteristics of urban green spaces.
5) The application potential of emerging data sources, such as street view imagery, in urban green space function assessment and pattern characterization, and the development of data-driven methods for fine-grained urban green space management.

By addressing these research gaps and goals, this research aims to contribute to the development of more comprehensive, nuanced, and actionable approaches to understanding and managing urban green spaces. The proposed innovations and methodological advancements are expected to bridge the gaps between different research domains and stakeholder groups, fostering a more human-centered understanding of urban greenery and facilitating the engagement of diverse stakeholders in envisioning and shaping sustainable urban futures.

The Chapter 3 will introduce a comprehensive assessment framework for urban green spaces that integrates multi-source data and evaluation metrics (GVI, GCR) to assess the spatial distribution, visibility, and composition of urban green spaces in a multi-dimensional manner. This framework will provide a foundation for the subsequent chapters, which will focus on the development of the S3PVI indicator and the application of advanced visualization technologies in urban green space assessment and modeling. By integrating quantitative measures with aesthetic considerations and capturing the temporal dynamics of urban vegetation, this research aims to provide urban planners and designers with evidence-based tools and insights to guide plant selection and configuration decisions, ultimately creating visually appealing, engaging, and sustainable urban green spaces.

Chapter 3

Comprehensive assessment framework for urban green spaces

3.1 Overview

Despite the growing recognition of the importance of urban green spaces in promoting environmental, social, and economic sustainability in cities, current methods for assessing urban green spaces often focus on single dimensions. While these indicators quantify the state of urban green spaces from certain perspectives, they fail to fully capture the multifunctionality and internal heterogeneity of these spaces. To overcome the limitations of existing approaches, this chapter introduces an innovative comprehensive assessment framework that integrates multi-source data and advanced technologies to characterize the complex features of urban green spaces from multiple dimensions, including spatial distribution, visibility, and composition. This framework cleverly combines GCR, GVI, and image classification techniques, aiming to provide urban planners and decision-makers with more comprehensive, detailed, and actionable evidence-based decision support.

To test the effectiveness and practicality of the proposed assessment framework, Osaka City, Japan, was selected as a case study. As a densely populated urban center, Osaka boasts diverse types of green spaces, such as parks, street greenery, and waterfront areas (*Osaka Prefectural Government,* 2020), while facing numerous challenges brought about by the urbanization process, such as insufficient green space and uneven distribution (Haaland & Van Den Bosch, 2015). Therefore, it serves as an ideal subject for validating the framework. Through empirical analysis in Osaka, this study aims to unveil the spatial pattern characteristics and influencing factors of the city's urban green spaces, as well as demonstrate the versatility and potential of the proposed assessment framework in guiding green space planning and management in different urban contexts.

Specifically, the main objectives of this chapter include:

1) Proposing a comprehensive assessment framework that integrates GVI, GCR, and image classification techniques to quantitatively characterize the multidimensional features of urban green spaces.

2) Applying the framework to Osaka City to systematically analyze the spatial distribution, visibility, composition, and quality of its urban green spaces.

3) Validating the accuracy and reliability of the assessment results from multiple perspectives through field surveys and statistical analyses.

4) Based on the assessment results, discussing the implications for urban green space planning practices in Osaka and other cities, providing new ideas for future urban green space research and management.

Chapter 3.2 will introduce the construction process of the comprehensive assessment framework in detail, including indicator selection, data processing, and analysis methods. Then, the study area and the multi-source data used will be described. In the results section, the multidimensional assessment results of Osaka's urban green spaces will be presented, focusing on the relationship between green space distribution and factors such as topography and land use. Finally, in the discussion section, the limitations of the study will be examined, and ideas for further improving the assessment framework and expanding its application scenarios in the future will be proposed.

3.2 Method and materials

3.2.1 Research framework

This study proposes a comprehensive assessment framework that integrates multi-source data and advanced technologies to evaluate the complex characteristics of urban green spaces. Figure 3.1 illustrates the overall framework of this research.

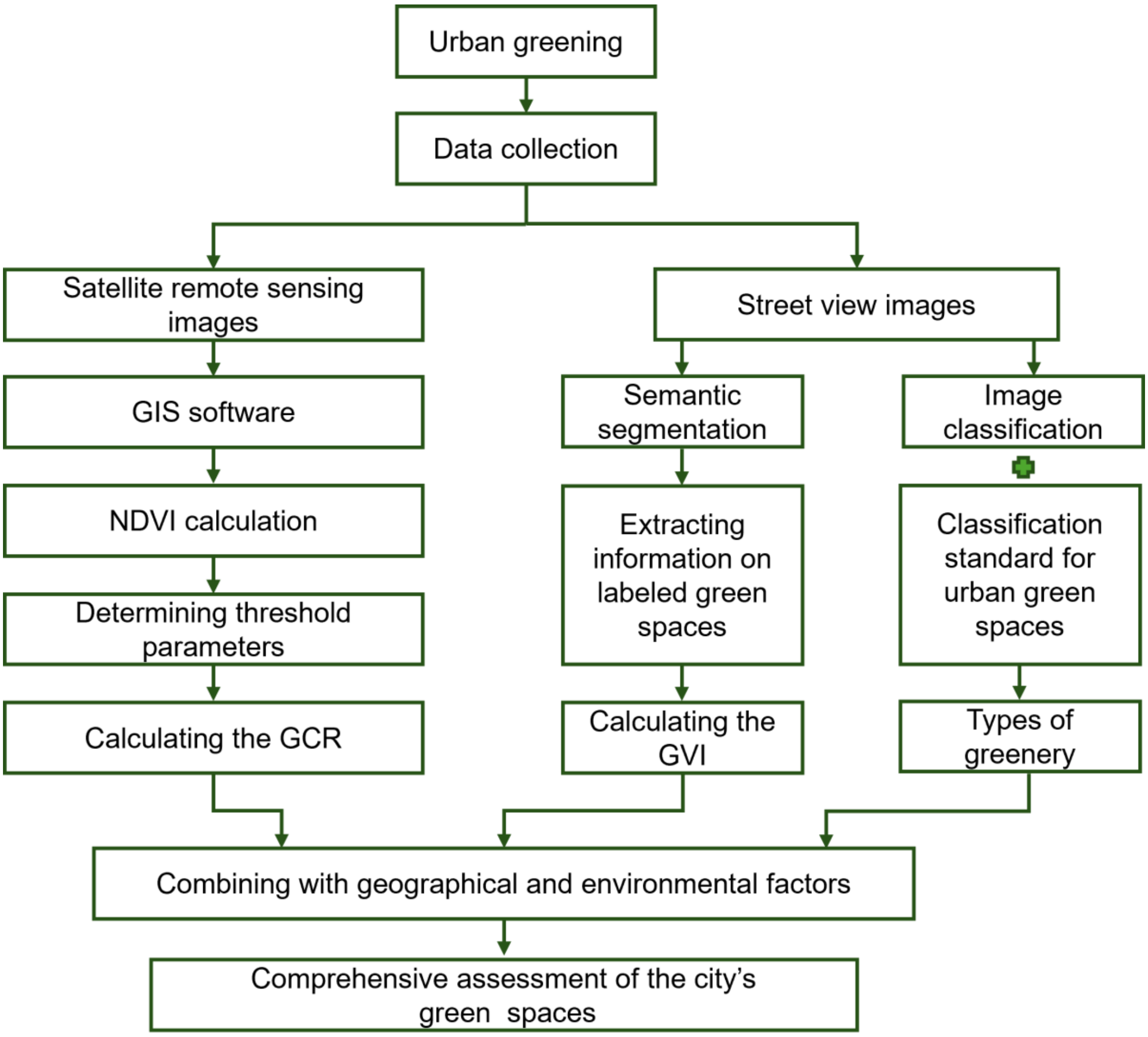


Figure 3.1 Workflow of the urban green space assessment framework.

Firstly, Sentinel-2 satellite remote sensing images are atmospherically corrected and processed for NDVI calculation using geographic information system (GIS) software. By comparing with field survey data, the optimal threshold parameters are determined for extracting green space information and calculating the GCR of each block or administrative unit. GCR, as an essential indicator of urban green space level, quantitatively reflects the proportion of green space in the two-dimensional plane.

Next, street view images are collected in bulk through the Google Maps API, and the PSPNet semantic segmentation model is employed to process these images and extract green space information for each pixel. PSPNet is a deep learning model based on pyramid scene parsing, which achieves fine-grained semantic segmentation of street view images by combining local and global contextual information. Based on the segmentation results, the GVI is calculated for each sampling point. GVI reflects the visibility of green spaces from a pedestrian perspective and is a crucial indicator for evaluating the visual effect of urban green space landscapes.

After obtaining GCR and GVI data, the EfficientNet image classification model is utilized to categorize street view images into different green space types, such as park green spaces, road green spaces, and riverside green spaces. The classification of green space types refers to national and industry standards.

Finally, by overlaying and analyzing multi-dimensional data such as GCR, GVI, and green space types, combined with geographical and environmental factors like urban topography and land use, the spatial pattern of urban green spaces and its influencing factors are systematically assessed, providing decision-making references for urban green space planning and management.

Sections 3.2.2 to 3.2.4 will introduce in detail the data sources used in the study and their preprocessing methods, the calculation process of GCR and GVI, and the classification method of green space types.

3.2.2 GCR calculation

The GCR was calculated using the Sentinel-2 satellite imagery and the NDVI. The NDVI is a widely used remote sensing index that measures the greenness of vegetation based on the reflectance of red and near-infrared light (Carlson & Ripley, 1997). The NDVI ranges from -1 to 1, with higher values indicating denser vegetation cover.

The process of determining the GCR of a designated city involves pre-processing satellite imagery using the Sen2Cor algorithm for atmospheric correction and generating the NDVI. The NDVI data is then utilized in conjunction with GIS software to establish accurate measurements. Random ground-truth points are generated within GIS software to serve as reference locations for GCR calculations, ensuring representative coverage across the study area.

To accurately differentiate between vegetated and non-vegetated areas, a systematic adjustment procedure is applied. This involves randomly selecting 200 test points within the study city using GIS geoprocessing tools and visually inspecting them to iteratively adjust the NDVI thresholds. The most suitable thresholds are identified based on careful analysis, considering the specific context of the study and the influence of factors such as city and season on the vegetation threshold.

Once the appropriate thresholds are determined, the GCR is calculated for each sub-area in the designated study city using the regional statistics function of the remotely sensed normalized NDVI data. This calculation is performed through the spatial analysis tool provided by the GIS software. By adhering to this rigorous method, the GCR values are accurately determined, providing a quantifiable measure of the degree of green cover in each sub-area of the study city. The GCR for each administrative unit within the target area is calculated by counting pixels classified as vegetated and dividing by the total number of pixels， using the Equation (1):

$$GCR = \frac{Number\ of\ vegetated\ pixels}{total\ number\ of\ pixels} \times 100\% \quad (1)$$

where the number of vegetated pixels is the count of pixels with NDVI values greater than or equal to the vegetative threshold, and the total number of pixels is the count of all pixels within the administrative unit.

3.2.3 GVI calculation

The GVI was calculated using Google Street View imagery and a deep learning-based semantic segmentation model, PSPNet (Zhao et al., 2017). The model was pre-trained on the Cityscapes dataset (Cordts et al., 2016) and fine-tuned on a custom dataset of street-level imagery from the target city. After inputting the image, the annotated image can be obtained as shown on the right side of Figure 3.2. The input image is first passed through an encoder network to extract features at different scales. Then, the pyramid pooling module (PPM) is applied to capture context information at multiple scales. The PPM outputs are concatenated with the encoder features and fed into a decoder network to generate the final pixel-wise predictions. The PSPNet model can effectively capture both local and global context information, enabling accurate semantic segmentation of urban green spaces.

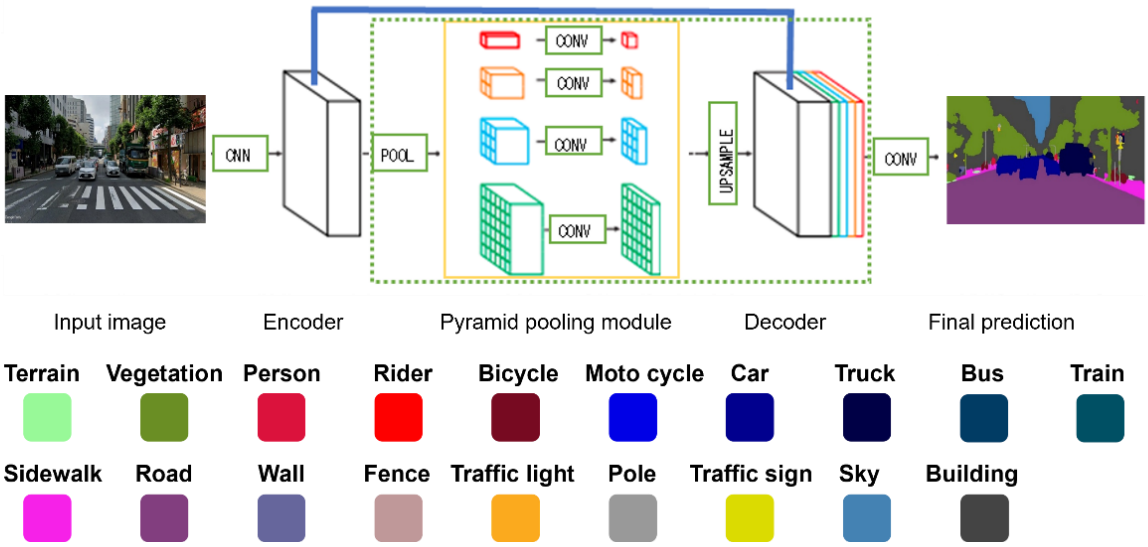


Figure 3.2 Schematic diagram of the PSPNet image segmentation model.

Google Street View images and GVI measurements have become popular in urban greening studies as they provide a pedestrian's perspective. The OSMnx library in Python is used to retrieve coordinate data of intersections within the study city from OSM. The GSV static API is then used to download street-level imagery by specifying parameters such as size, location, heading, field of view, pitch, and developer's key. GSV images (size = 640 × 640) are downloaded and stored for GVI calculations, excluding any inaccessible files due to privacy restrictions, coverage limitations, or other issues. The downloaded images are then filtered before being used for research. For a comprehensive assessment, all intersections within the road network of the target area are selected as sample points. At each intersection, images are captured from multiple directions (i.e., 0°, 60°, 120°, 180°, 240°, and 300°) to ensure a complete panoramic view, as shown in Figure 3.3. These multi-angle images provide a comprehensive view of the urban green space at each location, enabling the calculation of the GVI and other visual indicators. The PSPNet model analyzes these images to quantify the visible greenery, calculating the GVI based on the proportion of green pixels to the total number of pixels in the images.

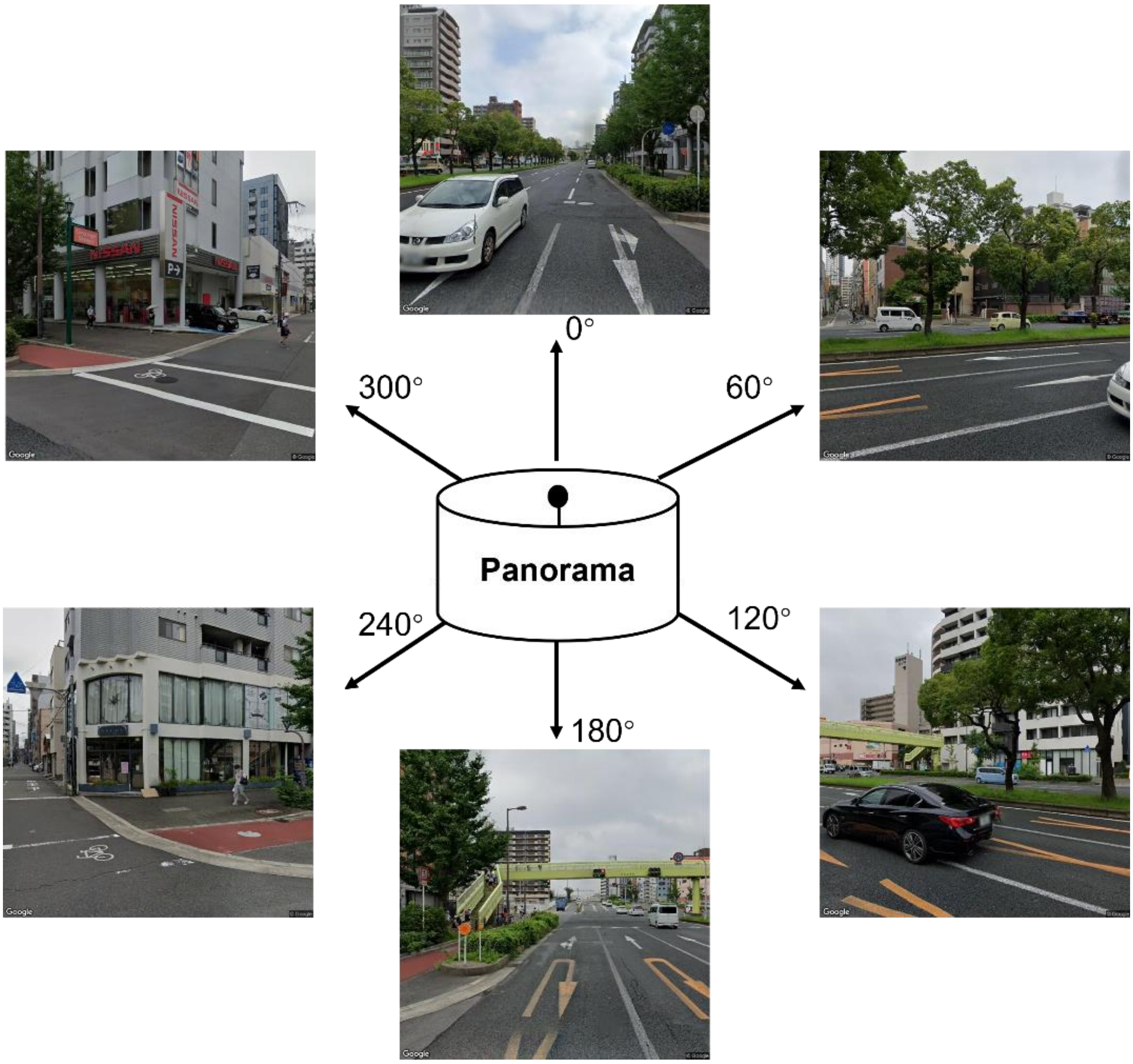


Figure 3.3 GSV images captured at a sample site in the study area from six different directions (0°, 60°, 120°, 180°, 240°, and 300°).

The PSPNet model was then applied to each image to extract the vegetation pixels and calculate the GVI using the Equation (2):

$$GVI = \frac{\sum_{i=1}^{m} Area_{g_i}}{\sum_{i=1}^{m} Area_{t_i}} \times 100\% \qquad (2)$$

where m = 6, $Area_{gi}$ represents the total number of green pixels at the intersection in an image taken horizontally in the direction i = 1 to 6. In contrast, $Area_{ti}$ represents the total number of pixels in the image. The number of vegetation pixels is the count of pixels classified as vegetation by the PSPNet model, and the total number of pixels is the count of all pixels in the image.

3.2.4 Green space type classification

To classify green space types in the target area, Google Street View imagery and the EfficientNet image classification model (Tan & Le, 2020) were utilized. EfficientNet is an advanced convolutional neural network (CNN) known for its high accuracy and efficiency across various image classification tasks.

The dataset for this classification comprises manually categorized street-level images reflecting various green space categories such as parks, temples, rivers, and residential greening, identified based on urban landscape and policy guidelines relevant to the region.

The EfficientNet model, pre-trained on the ImageNet dataset (Deng et al., 2009), was then fine-tuned on the custom dataset for 50 epochs with a batch size of 32 and a learning rate of 0.0001. The model achieved an overall accuracy of 95% on the testing set.

The fine-tuned EfficientNet model was subsequently applied to classify the green space types in each administrative unit using the remaining street-level images. The proportion of each green space type was calculated by comparing the number of images classified as a specific type to the total number of images within the administrative unit.

3.3 Study area and data sources

3.3.1 Study area

Osaka, a prominent urban center situated in the Kansai region of Honshu, Japan's main island, presents a unique and compelling case for studying urban greening efforts (Figure 3.4). The city's rich historical background, diverse population density, and distinctive urban landscape provide an intriguing context to explore the dynamics of urban greening (Urban Area of Osaka City, 1955).

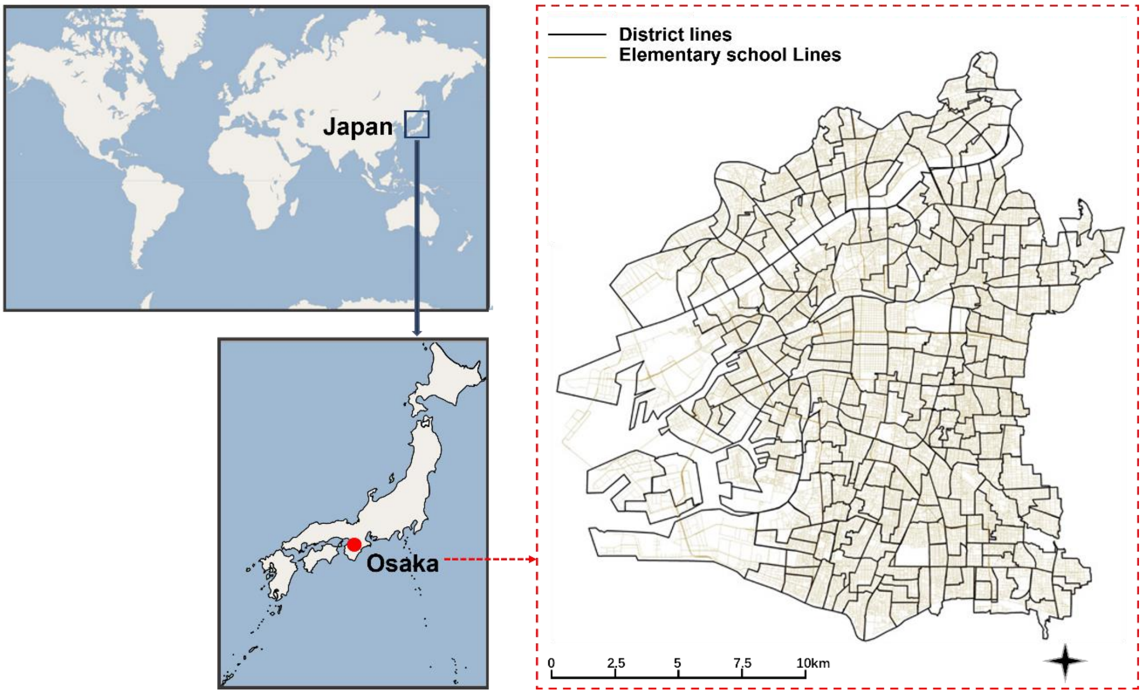


Figure 3.4 Geographical location of the study area.

Osaka's advantageous location along the mouth of the Yodo River has made it a vital port town and a strategic distribution hub for western Japan. The city has expanded to cover an area of 223 square kilometers, encompassing 24 administrative and 284 elementary school districts (List of Distinctive Town Names in Osaka City, 2022).

With a population of 2.7 million residents (Osaka City: 2020 Population Census Results, 2020), Osaka is Japan's third most populous city, with a population density of approximately 12,108 people per square kilometer, making it one of the most densely populated urban areas globally.

Osaka City's geographical formation primarily consists of an alluvial plain resulting from sediment accumulation of the Yodo and Yamato Rivers, naturally lacking in significant greenery. However, urban planning initiatives have been consistently geared towards enhancing both the quality and quantity of greenery through the establishment of new green spaces (Basic Plan of Greening in Osaka City, 2013).

Despite the progress made towards these objectives, a considerable discrepancy persists between the status and the targeted goals set for the mid-21st century. Given the slight but notable decline in available green space (-1.13%) over time and the escalating importance of accessible green areas (C. Huang et al., 2017), understanding the distribution and accessibility of green spaces within Osaka City has emerged as a critical research imperative.

This study, while underscoring its methodological generalizability, utilizes Osaka as a case study for practical application. Chosen for its unique geographical, historical, and demographic attributes alongside data availability, Osaka's insights can enhance broader knowledge and guide urban greening initiatives in comparable urban settings.

3.3.2 Data sources

To assess the urban green spaces in Osaka City, multiple data sources were employed, including Google Street View imagery, remote sensing satellite imagery, and geospatial data from various public sources. Table 3.1 summarizes the main data sources utilized in this study, detailing the parameters for accessing and downloading the data.

Table 3.1 Data sources used in the study

Data type	Source	Acquisition year
Sentinel-2 satellite imagery	Multispectral Instrument, Level-2A (Sentinel-2 MSI)	2021
Google Street View imagery	Google Web Service API	2019
Geographic data information	MLIT (Ministry of Land, Infrastructure, Transport and Tourism)	2008
School district boundaries	Gaccom	2022
OpenStreetMap data	OSMnx Python Library	2021

Street-level imagery was accessed through the Google Street View static API, allowing retrieval of images by specifying parameters such as size, location, heading, field of view, and pitch. The imagery was downloaded to perform GVI calculations, and the parameters were tailored to ensure the horizontal view mimics that of a pedestrian.

The use of satellite imagery from Sentinel-2 provided high-resolution data necessary for remote sensing analysis, captured on October 2, 2021. This imagery helped calculate the GCR of Osaka City.

For the geographical context, administrative boundaries were derived from MLIT datasets, and school district information, which delineates the primary unit of analysis, was obtained from Gaccom, offering a detailed breakdown of urban structure relevant to daily resident experiences.

The OSM data was crucial for obtaining detailed street layouts and was processed using the OSMnx library, which facilitates the extraction of comprehensive urban network data and geospatial analysis.

This combination of data sources ensures a comprehensive assessment of Osaka City's green spaces, integrating perspectives from both ground-level and aerial views to provide a robust framework for urban greenery analysis.

3.4 Results

3.4.1 Data integration results

The calculation of the GCR involved a systematic approach using NDVI data and ArcGIS map geodata software. The analysis involved a systematic approach to collect data and calculate the GCR, as depicted in Figure 3.5, which illustrates the NDVI data for Osaka.

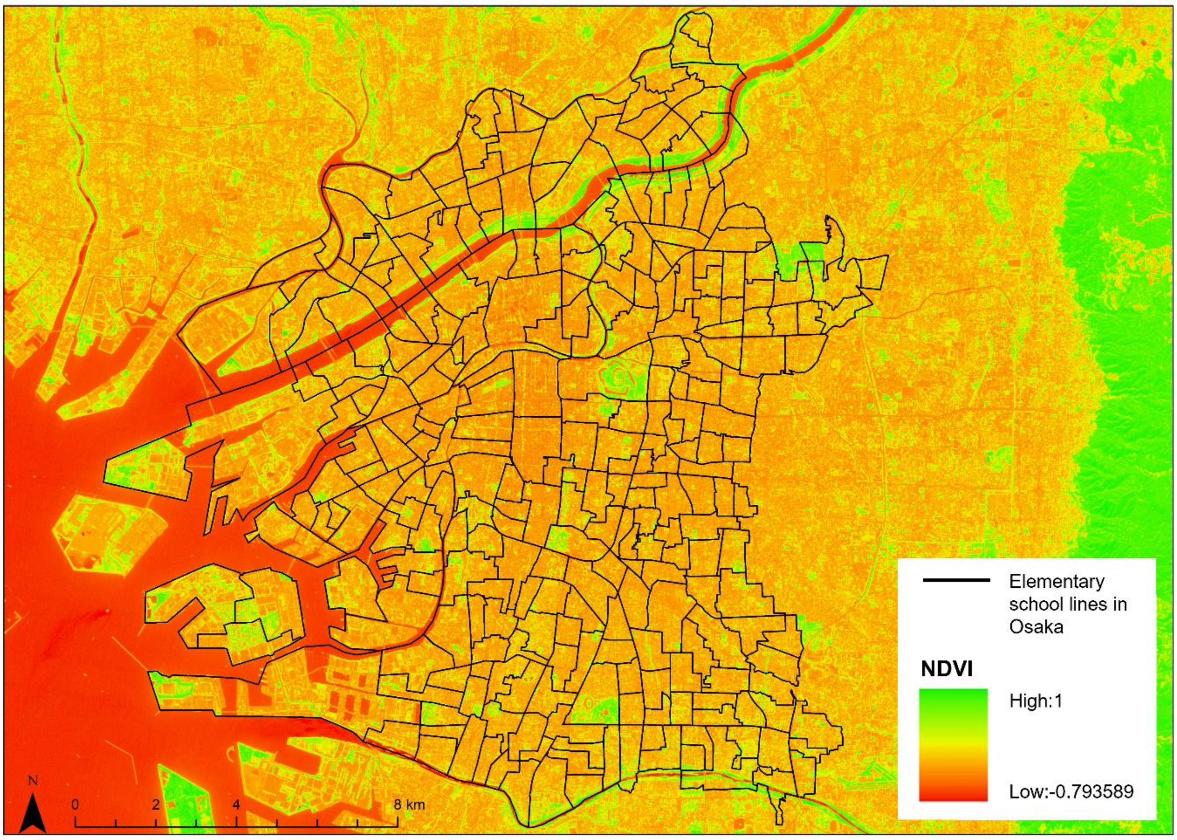


Figure 3.5 Distribution of NDVI in Osaka city.

To determine accurate measurements, a methodical adjustment process, i.e., described in section 3.2.2 was adopted to determine thresholds that distinguish vegetated and non-vegetated areas in Osaka City. The data from Osaka City were rigorously visually checked and iteratively adjusted to determine the NDVI threshold value of 0.274, which is most suitable for the study environment. Using this threshold, the GCR of each elementary school district in Osaka was calculated using a regional statistical function of remotely sensed normalized NDVI data. The resulting GCR values accurately quantified the extent of green coverage in each elementary school district in Osaka.

The GVI calculation involved the collection of GSV images using the Google Street View image API as described in Section and subsequent image processing as described in Section 3.2.3. A total of 273,462 GSV images of Osaka City were successfully used for the GVI calculation. These images were semantically segmented using the PSPNet model, which resulted in the accurate labeling of the images into 19 categories. The labeled images were then used to calculate the GVI for each location based on the GVI Equation (Equation 2). The average GVI value for each location was calculated, and the GVI distribution for Osaka was visualized using geographic data analysis in ArcGIS Map. The GVI was divided into five classes (<2.5%, 2.5% - 5.0%, 5.0% - 7.5%, 7.5% - 10.0%, and >10%) based on the average GVI value of 7.46 in Osaka.

To provide a more precise depiction of the GVI distribution, the example of the Sennba area in Chuo-ku, Osaka City, was chosen, and the average GVI degrees for each road intersection within this area are illustrated on the right side of Figure 3.6. Sennba refers to a specific district in Chuo-ku, Osaka City, Osaka Prefecture, known as the central business district of Osaka City.

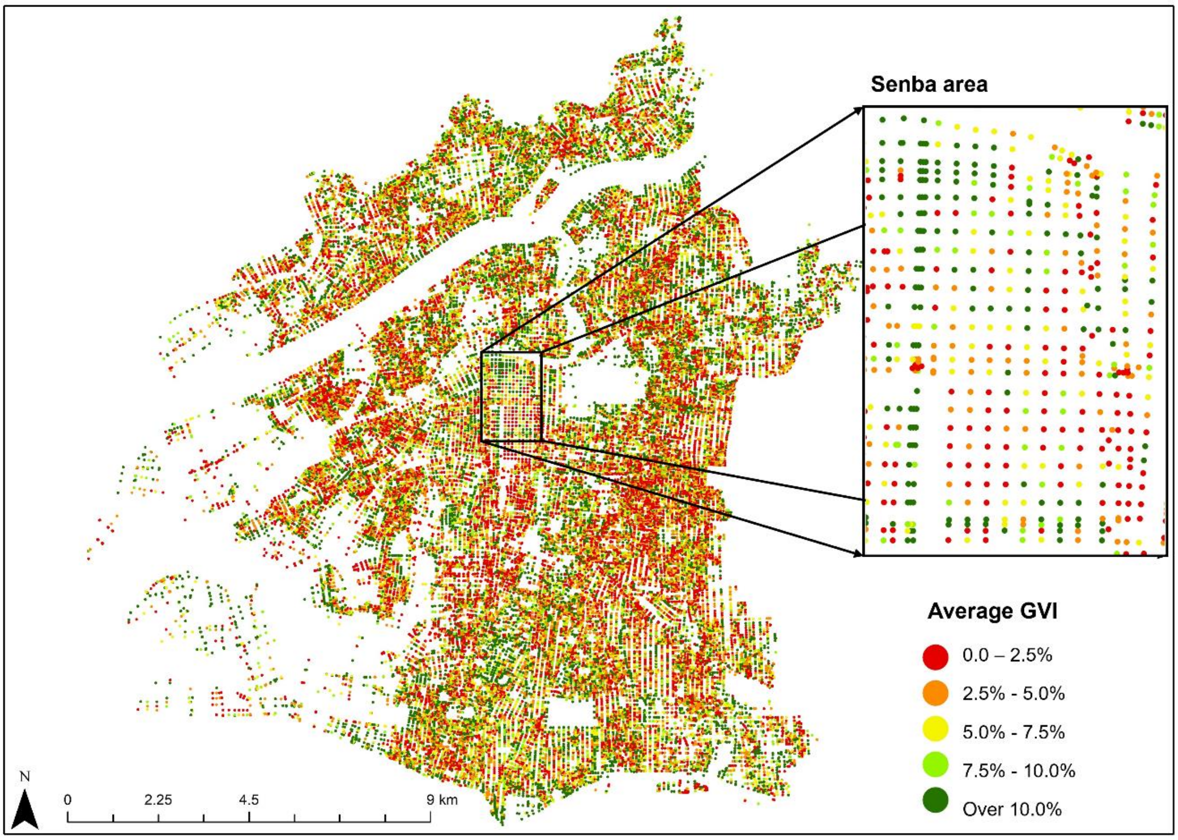


Figure 3.6 Spatial distribution of average GVI in Osaka.

By integrating GCR and GVI data, a comprehensive assessment of urban greening in Osaka City was conducted. Based on the average GVI and GCR values, 284 elementary schools were categorized into four different groups, as shown in Figure 3.7. This categorization highlights the different levels of green space coverage and visibility throughout the city.

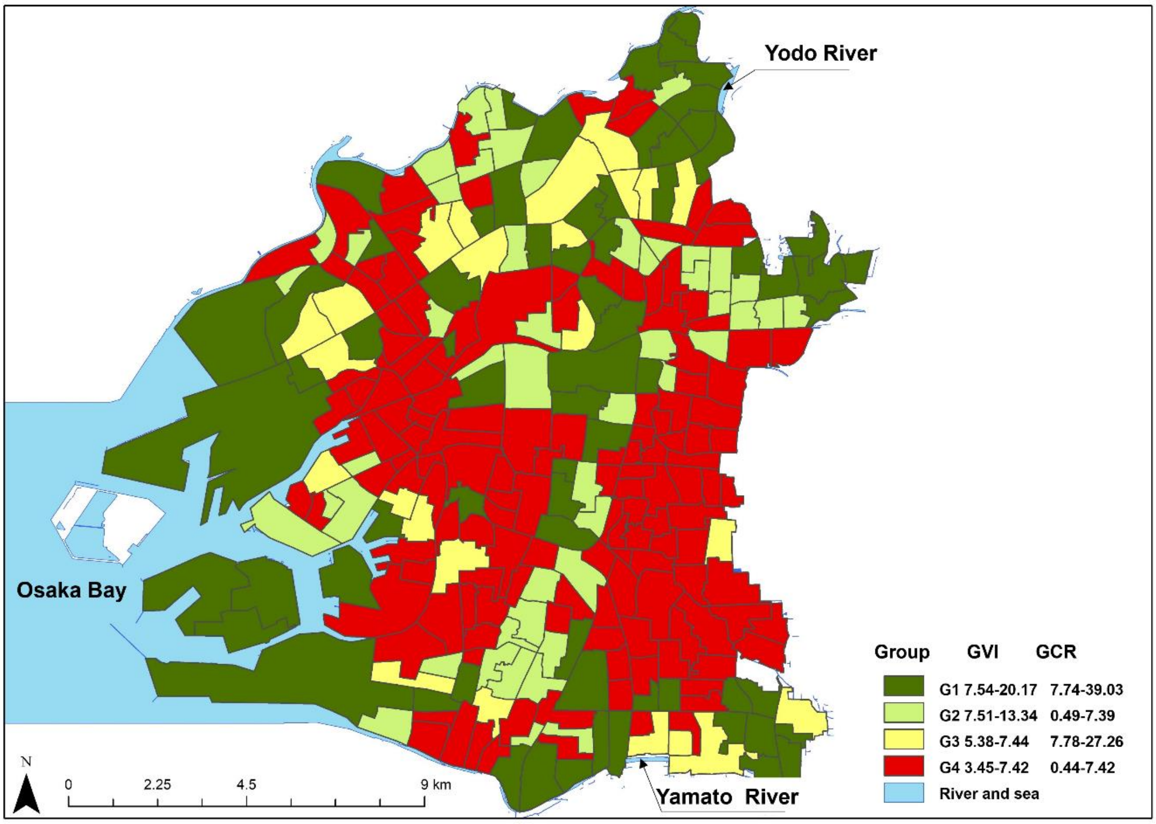
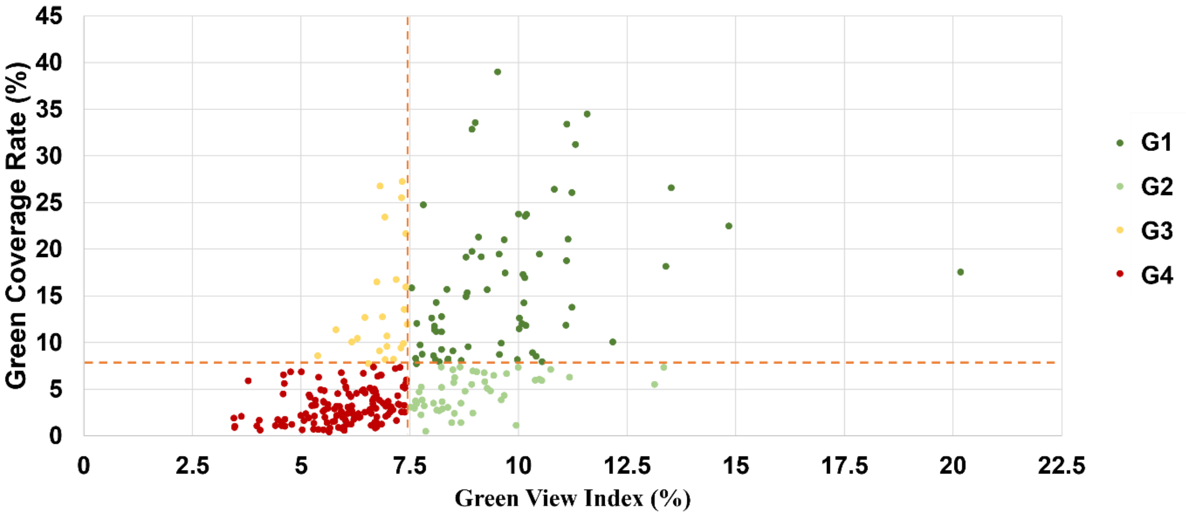


Figure 3.7 Comparison of GVI and GCR in Osaka's elementary school districts.

As shown in Figure 3.8, Group 1 includes 69 elementary school districts with superior green visibility and coverage. The average GVI is 9.67%, reaching up to 20.17%, while the average GCR stands at 16.14%, with a maximum of 39.03%. Schools such as Tomobuchi and Nanko showcase high GVI values, whereas Yakino and Taishibashi excel in GCR, indicating a robust green presence both visually and spatially.



	Number of schools	Percentage (%)	Average GVI (%)	Average GCR (%)
G1	69	24.30	9.67	16.14
G2	49	17.25	8.95	4.63
G3	24	8.45	6.87	14.10
G4	142	50.00	5.97	3.12
SUM	284	100.00		

Figure 3.8 Relationship between GVI and GCR in four groups of elementary schools in Osaka.

Group 2 consists of 49 districts where green visibility outstrips coverage. It has an average GVI of 8.95% and a maximum of 7.51%, coupled with a relatively low GCR average of 4.63% and a maximum of 7.49%. This disparity suggests a visually green environment that lacks substantial physical greenery.

Group 3 encompasses 24 districts characterized by adequate green coverage but lesser visibility. Here, the average GVI is lower at 6.87%, with a maximum of 7.44%, whereas the GCR is more favorable, averaging 14.10% with a peak at 7.78%. This configuration implies a scenario where green spaces are prevalent but perhaps not as prominently visible.

Group 4, the largest group with 142 districts, shows deficiencies in both green visibility and coverage. The average GVI is only 5.97%, dipping to a low of 3.45%, and the GCR averages at a meager 3.12%, with the minimum at 0.44%. This indicates areas with sparse green spaces, both in terms of visibility and actual coverage.

3.4.2 *Image classification results*

A dataset of 1,970 street-level photos in Osaka was classified using the EfficientNet-b7 model, which achieved a high accuracy rate. The photos were classified into 10 types of green spaces based on the guidelines of the Ministry of Land, Infrastructure, Transport and Tourism (Ministry of Land, Infrastructure, Transport and Tourism, 2008), including photos of parks (n=205), temples and shrines (n=202), rivers (n=201), agricultural land (n=109), vacant land (n=201), street greenery (n=230), residential greenery (n=287), commercial greenery (n=205), industrial greenery (n=130), and other (n=200).

This classification was instrumental in analyzing the distribution and types of green spaces in the 273,462 GSV images collected for this study. 70% of the training dataset was used for training, 15% for validation, and 15% for testing, resulting in a final accuracy of 0.941 after training EfficientNet-b7. The results, shown in Figure 3.9, provide a detailed overview of the spatial distribution and types of green spaces in Osaka City, highlighting the critical role of diverse green spaces in urban sustainability. This detailed and accurate analysis provides a comprehensive overview of the spatial distribution, visibility, and type of green spaces in Osaka City, facilitating targeted urban green space planning and management strategies.

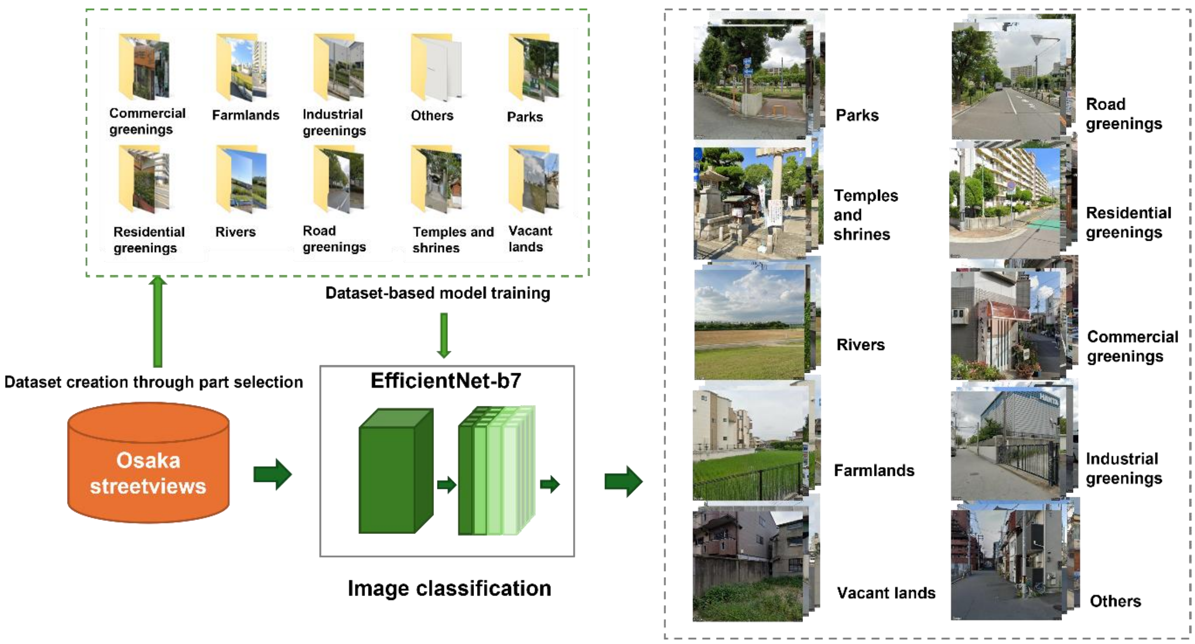


Figure 3.9 Classification of green space types in Osaka based on GVI.

3.4.3 Correlation between topography and green space

Analyzing the relationship between topography and green space is crucial because topography significantly influences the distribution and characteristics of green spaces in urban environments. Topographic features such as elevation, slope, and proximity to water bodies can dramatically affect the availability and suitability of land for green spaces. This analysis aims to understand how natural and man-made landscape features collectively determine the spatial patterns of green spaces in urban settings.

As depicted in Figure 3.10, Osaka City is intersected by the Yodo River to the north and the Yamato River running east-west in the south. The city's central region is marked by the medium and low terraces of the Uemachi Plateau, extending north-south. The landscape west of the Uemachi Plateau to the bay area predominantly consists of reclaimed land, including deltaic lowlands, natural levees, sandbars, and dunes. Conversely, the area east of the Uemachi Plateau features fan lowlands, natural levees, sandbars, dunes, and delta lowlands, with natural levees, sandbars, and dunes dominating the north of the Plateau.

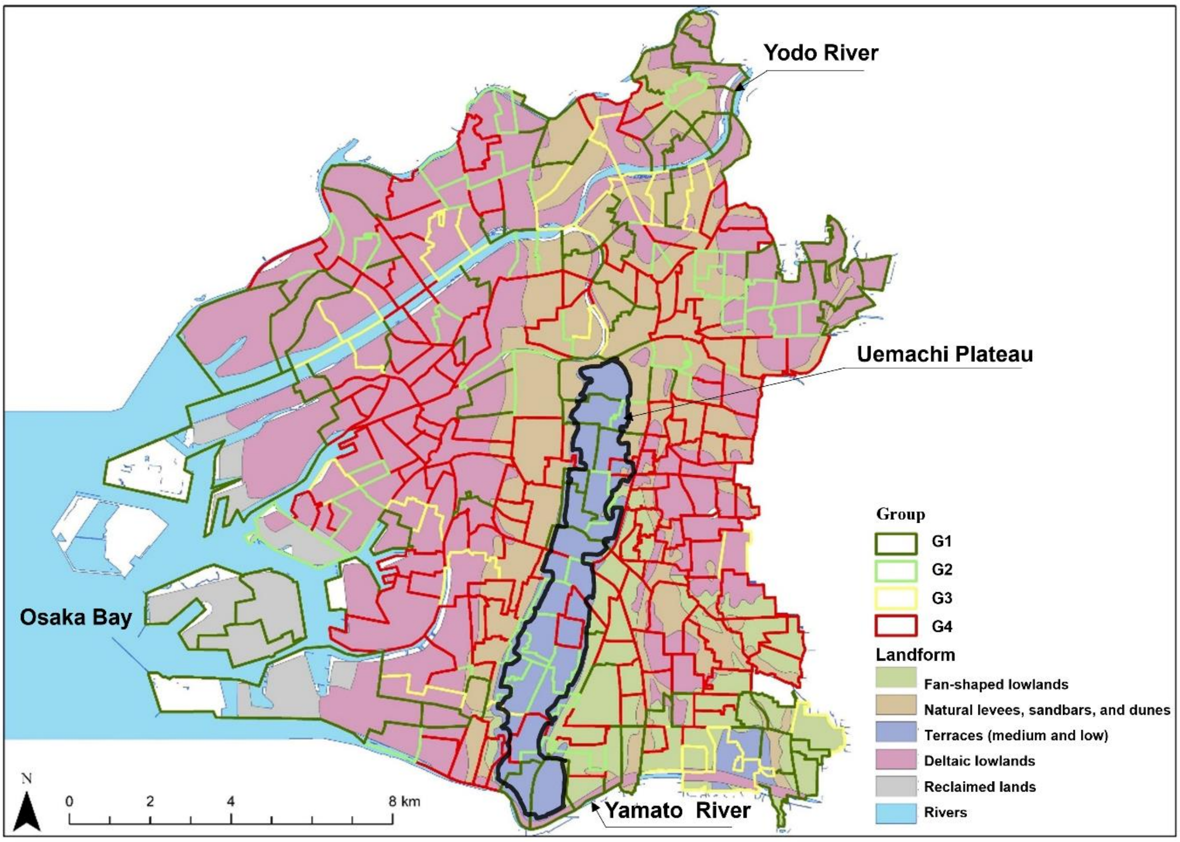


Figure 3.10 Classification of elementary school districts based on terrain distribution.

Cross-analysis of these topographic classifications with the groups of elementary school districts reveals certain patterns. Group 4, characterized by low green visibility and cover, predominantly occupies the deltaic lowlands west and east of the Uemachi Plateau and in the northern part of the Plateau, mainly on natural levees, sandbars, and dunes. In contrast, Group 1, which exhibits high green visibility and cover, is primarily located on reclaimed land and deltaic lowlands in the bay area to the west and on the terraces of the Uemachi Plateau. This group is also distributed along the Yodo and Yamato Rivers, in the deltaic lowlands in the east, and in the fan-shaped lowlands in the southeast. Group 2, with high green cover but low visibility, is scattered across the Uemachi Plateau, the deltaic lowlands in the northeastern and northern city regions, and natural levees, sandbars, and dunes. Lastly, Group 3, marked by high green visibility but low cover, is predominantly situated along the Yodo and Yamato Rivers.

3.4.4 Data on green space in four groups of elementary schools

This study analyzed the GVI values of different types of green spaces in Osaka City to investigate the distribution of green spaces.

Figure 3.11 shows the average GVI values for each type of green space. Parks have the highest average GVI at 39.62%, followed by temples and shrines with an average GVI of 21.97%. The average GVI values for road greening, vacant land, and residential greening are 19.35%, 17.65%, and 17.50%, respectively. Rivers have an average GVI of 14.86%, while industrial greening, commercial greening, and farmland have average GVI values of 12.55%, 9.69%, and 7.90%, respectively.

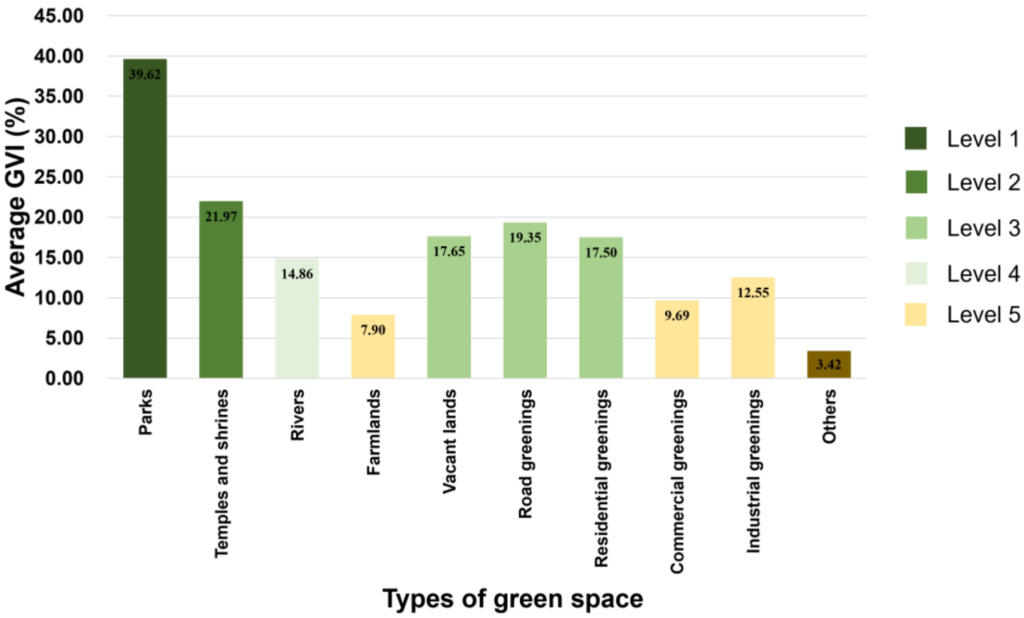


Figure 3.11 Average GVI values for different green space types in Osaka.

Figure 3.12, based on GSV images, comprehensively illustrates the proportion of green spaces in the four groups of elementary schools in Osaka City. By analyzing the proportions of each type of green space in Osaka City as a whole (SUM), the results are: "other" green spaces have the highest proportion at 54.99%, followed by residential greening (17.86%), commercial greening (9.45%), road greening (8.06%), industrial greening (3.12%), parks (2.22%), vacant land (1.83%), temples and shrines (0.92%), farmland (0.88%), and rivers (0.69%).

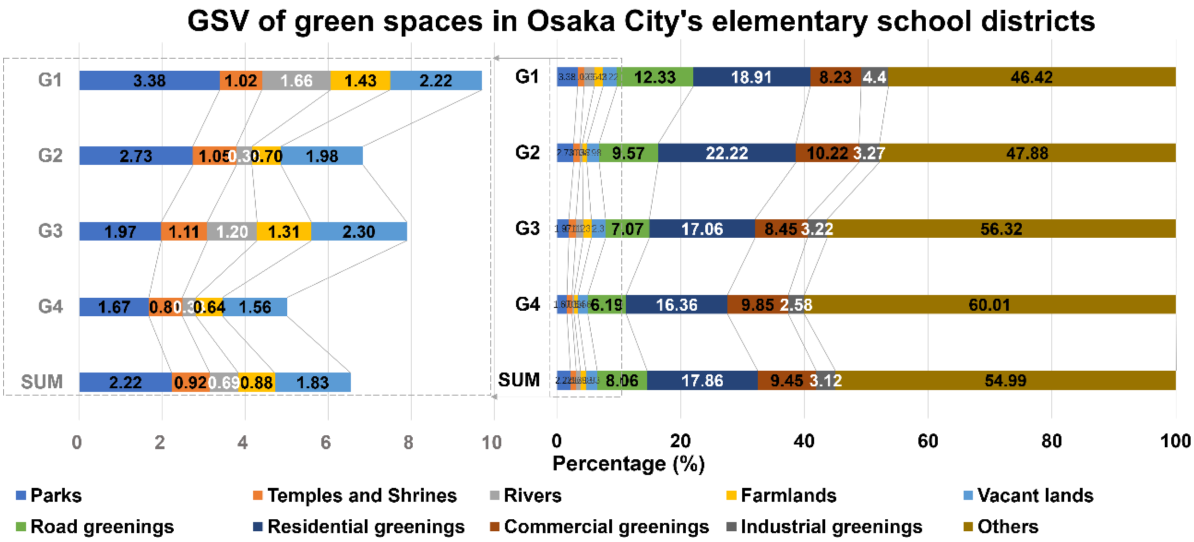


Figure 3.12 Proportion of different green space types in Osaka's elementary school districts based on GSV image analysis.

All groups have a high percentage of "other" green spaces. Compared to the entire city (SUM), Group 1 has higher proportions of green spaces in categories such as residential greening, road greening, parks, temples, and shrines. Group 2 also has higher proportions of green spaces in residential greening, commercial greening, road greening, parks, temples, and shrines compared to the city average.

3.5 Discussion

3.5.1 Distribution and characteristics of green space types

This study conducted a comparative analysis of green space data in Osaka, revealing the distribution and characteristics of various types of green spaces. By using GVI as a quantitative measure, the study provides an in-depth evaluation of the urban contributions of these green spaces. The results indicate that parks play a critical role in Osaka's green spaces, offering not only visual accessibility but also serving as important community gathering points and urban oases. Additionally, temples and shrines with high GVI values significantly enhance the city's green landscape, providing visual appeal and tranquility, while also contributing to cultural heritage preservation.

Road greening, vacant land, and residential greening constitute significant components of the city's green environment. Initiatives such as road greening enhance the visual quality of urban road networks and their ecological value. Vacant land presents opportunities for new green spaces, contributing to the city's overall greenery. Residential greening improves micro-scale greenery and elevates living conditions.

Although industrial and commercial greening, as well as farmland, have lower GVI values, their contributions to the city's overall greenery are significant, offering the potential for harmonious integration of green spaces within urban functions. Understanding the dispersion and characteristics of various green spaces enables policymakers and urban planners to identify areas needing enhancement. By improving the visual quality and accessibility of lower GVI green spaces, such as industrial and commercial areas, more visually pleasing and inclusive urban environments can be created. Simultaneously, preserving and expanding high-GVI green spaces, such as parks and temples, ensures the continuity of visually attractive and socially significant green areas.

3.5.2 Topographical factors influencing urban greening

Additionally, the study analyzed the geographical distribution and topographical factors influencing urban greening. The results indicate a correlation between topography and green space in Osaka City. Higher green visibility and coverage are primarily found in reclaimed land, deltaic lowlands, terraces of the Uemachi Plateau, and along rivers such as the Yodo and Yamato Rivers. These areas have a greater concentration of green spaces due to their historical significance and long-standing tradition of incorporating greenery. The proximity to water bodies also enhances the availability of green spaces in these regions.

3.5.3 Green space distribution among elementary school groups

The evaluation of green space data among different elementary school groups provides crucial insights into the distribution and quality of Osaka's urban greenery. The analysis reveals that "other" green spaces (areas with little greenery) dominate across all school groups, highlighting the need for city-wide green space augmentation. Group 1 outperforms other groups in overall green space quantity, driven by residential and road greening and high-GVI spaces such as parks, temples, shrines, and vacant lands. In contrast, Group 2 has a significant proportion of residential, commercial, and road greening, indicating these sectors' substantial role in shaping Osaka's green landscape.

3.5.4 Limitations and future research directions

Despite the integrated evaluation framework proposed in this chapter contributes significantly to the field, certain limitations remain when utilizing automatically downloaded street view images. As illustrated in Figure 3.3, the perspectives of the currently downloaded street view images differ from those of pedestrians in the real world. While using these data directly for analysis enables horizontal comparisons, it cannot be directly compared with photo data taken from a pedestrian's viewpoint. This discrepancy underscores the necessity for appropriate preprocessing and standardization of the data when employing street view images for urban green space assessment.

Furthermore, strict and consistent data collection protocols are imperative when calculating indicators such as the GVI. Parameters such as the aspect ratio and shooting distance of photos significantly influence the calculation results of these indicators. To ensure the reliability and comparability of assessment outcomes, future research should focus on refining data collection and processing procedures and establishing standardized data collection protocols.

This study also does not fully account for the differentiation of vegetation types and their spatiotemporal dynamics, lacking detailed characterization of vegetation type diversity and dynamic changes.

Furthermore, the research findings, primarily based on Osaka City, may not be universally applicable to urban environments with different geographic, socioeconomic, and ecological characteristics. The validity and reliability of results depend on the quality, resolution, and coverage of the data sources used. Cities with different urban morphologies, climate conditions, or cultural attitudes towards green spaces may exhibit significantly different patterns of green space distribution and usage. While the study focuses on the relationship between topography and green space, it may overlook other factors influencing urban greening, such as historical development patterns, local planning policies, or community initiatives. Additionally, the research provides a static snapshot at a specific point in time, potentially failing to account for future changes in urban green spaces. This limitation is particularly significant given the dynamic nature of urban environments and the increasing impacts of climate change on urban vegetation.

Expanding the research scope to consider a broader range of factors influencing urban greening, including socioeconomic variables, land use patterns, and local policies, would provide a more comprehensive understanding. For instance, examining the relationship between green space distribution and factors such as income levels, population density, or zoning regulations could reveal important insights into urban equity and environmental justice issues. Adopting dynamic research methods to reflect changes in urban green spaces over time is crucial, as these spaces are subject to transformation due to construction, redevelopment, and climate change. Long-term studies that track changes in green space over years or decades could provide valuable information on urban ecological trends and the effectiveness of green space policies.

These enhancements would contribute to a more holistic, adaptable, and accurate study of urban green spaces. By addressing the current limitations, future research can offer more valuable insights for urban planners and policymakers, facilitating optimized green space planning and management. This approach would not only unveil the multi-dimensional features of urban green spaces but also provide a more nuanced understanding of their roles and dynamics within diverse urban environments. Moreover, it could lead to the development of more sophisticated models for predicting the impacts of urban development on green spaces and guide the creation of more resilient and sustainable urban ecosystems.

3.6 Summary of this chapter

This chapter presents a comprehensive assessment framework that integrates multi-source data, including satellite imagery and street view images, with advanced technologies such as deep learning and GIS, to evaluate the spatial distribution, visibility, and composition of urban green spaces. The framework introduces a multi-dimensional approach to characterize urban green spaces, combining GCR, GVI, and image classification techniques.

The case study of Osaka City demonstrates the effectiveness of the proposed framework in revealing the heterogeneous distribution of green spaces and the influence of topographical factors on green space patterns. The analysis uncovers the distinct contributions of various green space types to the city's overall greenery and highlights the need for targeted strategies to enhance the visual quality and accessibility of lower GVI areas.

The research findings offer valuable insights for urban planners and policymakers to optimize green space planning and management. By identifying areas with limited green space provision or lower visual quality, the proposed framework can guide evidence-based decision-making and support the development of more equitable and sustainable urban environments.

However, the current study has certain limitations, such as the reliance on automatically downloaded street view images that may not fully represent pedestrian perspectives and the lack of detailed characterization of vegetation type diversity and temporal dynamics. These limitations underscore the need for future research to refine data collection protocols, explore the integration of multi-temporal data, and develop novel indicators to capture the spatiotemporal patterns of urban green space vegetation.

Chapter 4 will build upon the foundation laid in this chapter by introducing a multi-temporal visualization analysis framework and a new vegetation type scale visualization index, S3PVI. These advancements aim to provide a more dynamic and nuanced understanding of urban green space vegetation patterns, enabling refined green space planning and management.

In conclusion, the comprehensive assessment framework presented in this chapter contributes to the development of urban green space assessment and dynamic modeling techniques. The integration of multi-source data and advanced technologies opens up new possibilities for evaluating the multi-dimensional characteristics of urban green spaces.

Continued research and collaboration across disciplines are essential to address the limitations, enhance the robustness and applicability of the proposed methods, and ultimately support the creation of sustainable and resilient cities.

Chapter 4

A multi-temporal evaluation framework for the S3PVI

4.1 Overview

Chapter 3 proposes an integrated evaluation framework for urban green spaces that combines multi-source data and employs indicators such as GCR and GVI for a multidimensional assessment of urban green spaces. However, it does not yet address the analysis of vegetation types and their spatiotemporal differentiation characteristics. It lacks a detailed depiction of the diversity of vegetation types and their spatiotemporal dynamics, making it difficult to provide refined and dynamic references for green space planning and management.

To overcome the limitations of existing research that focuses only on greenness characteristics, this chapter constructs a multi-temporal visualization analysis framework for urban green space vegetation based on street view images. A new visualization index on the scale of vegetation types S3PVI is proposed to quantitatively describe the dynamic changes in the visibility of different plant types over time. Based on deep learning technology, an automated calculation method for the S3PVI index was developed, enabling large-scale, efficient extraction and analysis of vegetation information.

The framework proposed in this chapter is a decision support tool for urban planners, park designers, and street/road planners, rather than a simple planning or design evaluation tool. The framework employs 3D reconstruction technology to describe urban vegetation characteristics from multiple angles and at different times. Compared to traditional single-view assessments, this approach offers a more comprehensive and consistent evaluation method. By utilizing advanced deep learning models to extract and analyze vegetation information from modified street view imagery, it enhances both the efficiency and scalability of the analysis.

However, it should be noted that the current method has some limitations. Due to the small plant dataset and technical constraints, it is difficult to identify plants in non-ornamental seasons, such as cherry trees when not in bloom. This results in the statistics only identifying plants in their ornamental state. Nevertheless, even green plants that are not in their ornamental period are beneficial to humans, such as the year-round greening effect provided by evergreen plants and the warmth brought by deciduous plants not blocking sunlight in winter.

This chapter applies the proposed analysis framework to the Sanshikisaido area in Suita, Osaka Prefecture, Japan, revealing the spatiotemporal differentiation patterns of the visibility characteristics of different plant types in the study area. The visual contribution of landscape plants such as cherry blossoms and maple trees shows significant dynamic changes during the flowering season in spring and the foliage color change period in autumn. Further virtual scenario simulation analysis indicates that a reasonable vegetation configuration plan can effectively enhance the visual diversity and landscape rhythm of park green spaces across different seasons.

The research methods proposed in this chapter provide new technical support for the refined characterization of urban green space landscape features, assessment of vegetation configuration plans, and improvement of green space planning and management levels. Moreover, the spatiotemporal changes in vegetation visual characteristics revealed by multi-temporal analysis lay the foundation for the 4D vegetation landscape modeling based on NeRF and Stable Diffusion in Chapter 5.

4.2 Proposed framework

This section introduces the framework for multi-temporal urban green space vegetation visualization analysis, including key steps such as data collection, preprocessing, semantic segmentation, S3PVI calculation, and seasonal change analysis. Figure 4.1 shows the overall workflow of the method:

First, high-resolution street view images are collected. Then, the SfM algorithm is used to extract key points, including camera positions and angles, from multiple street view images. Using 3D Gaussian splatting technology, a smoother and more realistic 3D scene is reconstructed. Based on the 3D reconstructed scene, standardized views are generated at a fixed height (e.g., 1.6 m) and distance for observation.

It's important to note that when the original street view image data does not have significant differences and the angles cannot be adjusted, the standardized view generation step can be skipped, proceeding directly to plant visual feature evaluation based on S3PVI. This flexibility allows the framework to adapt to various data conditions and still produce meaningful results.

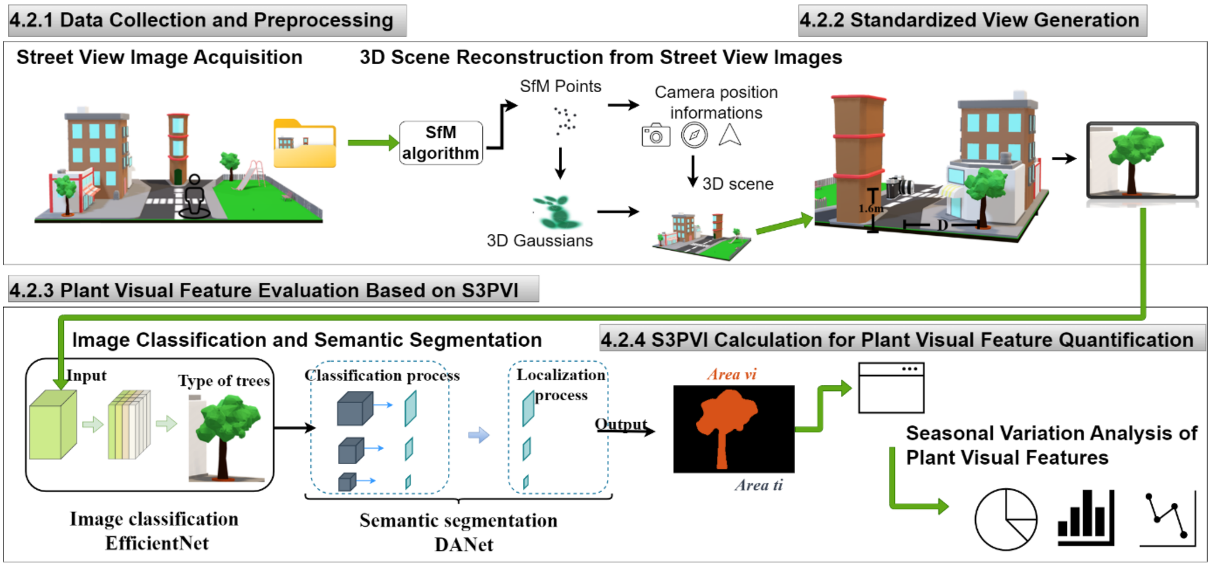


Figure 4.1 Workflow diagram of the multi-temporal urban green space vegetation visualization analysis framework.

In the image processing stage, the standardized view images are input into an EfficientNet-based classification network to identify tree species, and the DANet is used for semantic segmentation to determine the specific locations and ranges of plants (Tan & Le, 2020; Xue et al., 2019). The visible vegetation area ($Area_{v_i}$) and the total image area ($Area_{t_i}$) are calculated from the segmentation results to obtain the S3PVI index. The S3PVI index is then used to statistically analyze the seasonal changes in vegetation. This workflow demonstrates how to achieve detailed, multi-temporal analysis of urban green vegetation through street view images and advanced image processing technologies, combined with 3D reconstruction and standardized view generation.

4.2.1 *Data collection and preprocessing*

1) Street view image acquisition

The first step of the multi-temporal urban green space vegetation visualization framework is to collect street view images of the target urban area. These images can be obtained from online platforms such as Google Street View or Baidu Maps (Baidu Map Open Platform, n.d.; Google Earth, n.d.), ensuring that the collected data covers all angles of the scene and different times to capture the temporal changes in the visual appearance of the scene's plants.

2) 3D scene reconstruction from street view images using 3D Gaussian splatting.

To address the inconsistency issues in the original street view images, such as differences in camera angles and distortions, 3D Gaussian splatting technology is employed to reconstruct the 3D scene of the street view (Kerbl et al., 2023). This process includes the steps shown in Figure 4.2.

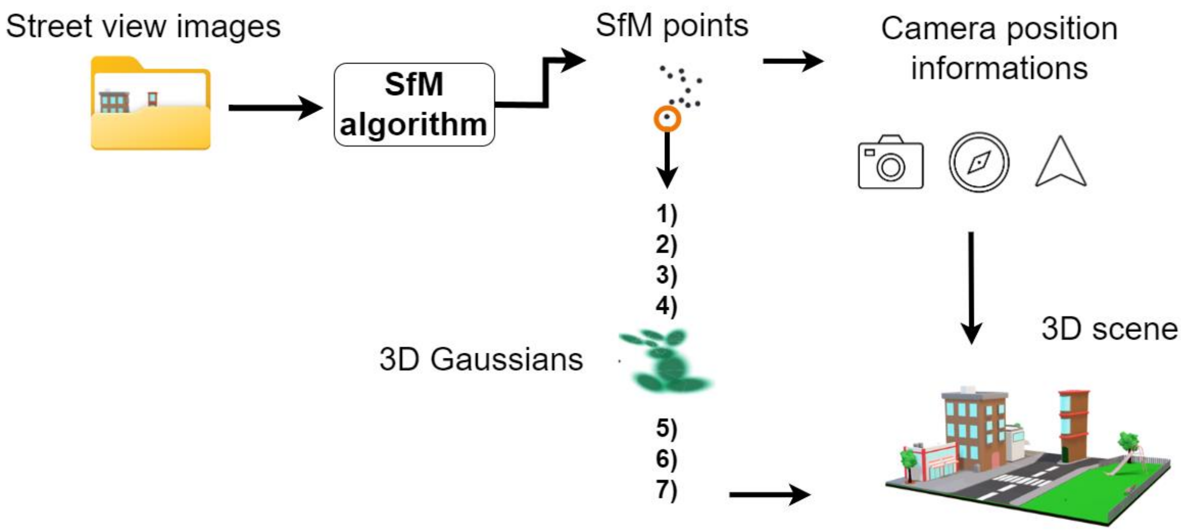


Figure 4.2 Process of 3D scene reconstruction using SfM and 3D Gaussian splatting

1) SfM: The SfM algorithm is used to reconstruct a 3D point cloud of the street view from two-dimensional street view images. The SfM algorithm estimates the positions and orientations of the cameras and the 3D coordinates of the scene points based on corresponding features identified in multiple images. The output of this step is a sparse 3D point cloud data.

2) 3D Gaussian splatting: The sparse 3D point cloud obtained from SfM is converted into a continuous volumetric density representation to obtain a smoother and more realistic 3D scene. This step utilizes the 3D Gaussian splatting technique, which is based on the principle of applying a Gaussian distribution weight to each point in the point cloud, converting the point cloud data into a continuous volumetric density representation.

Specifically, for each point $p_i(x, y, z)$ in the point cloud, a Gaussian kernel function is constructed as Equation (3):

$$G(p|p_i,\sigma) = exp\left(-\frac{\|p - p_i\|^2}{2\sigma^2}\right) \tag{3}$$

where $G(p|p_i,\sigma)$ represents the function value of the Gaussian kernel with center at point p_i and standard deviation σ at point p.Then, 3D Gaussian splatting is performed on the point cloud using these steps:

1) Initialize an empty voxel grid V to store the volumetric density.
2) Traverse each point p_i in the point cloud P.
3) For each point p_i, calculate its coordinates (x_v, y_v, z_v) in the voxel grid V.
4) Based on the Gaussian kernel function $G(p|p_i,\sigma)$, calculate the volumetric density increment of point p_i at the corresponding position (x_v, y_v, z_v) in the voxel grid V.
5) Accumulate the calculated volumetric density increment to the corresponding position in the voxel grid V.
6) Repeat steps 2-5 until all points have been traversed.
7) Perform thresholding on the voxel grid V to extract the 3D surface.

By superimposing the Gaussian kernel functions of all points, a continuous volumetric density representation is obtained, generating a smoother and more realistic 3D scene.

Finally, through the two steps of SfM and 3D Gaussian splatting, the reconstruction of a continuous 3D scene from 2D street view images is achieved. This method effectively addresses the inconsistency issues in the original street view images and lays the foundation for subsequent scene analysis and applications.

4.2.2 Standardized view generation

From the reconstructed 3D scene, standardized views are generated to simulate pedestrian perspectives and ensure consistency across different urban environments. The parameters in Table 4.1. are used to generate standardized views.

Table 4.1 Virtual camera parameter settings for standardized view generation

Parameter	Description
Camera height	The virtual camera is located at a height of 1.6 meters (Osaka Prefectural Government, 2020), representing the average eye level of pedestrians.
Camera distance	The distance between the virtual camera and the target plant is calculated using Equation (4) based on the plant height, ensuring full visibility of the plant.
Image aspect ratio	The aspect ratio of the generated views is 2:1, simulating the human field of view.
Plant positioning	The bottom edge of the plant trunk is aligned with the bottom of the image frame, and the top edge of the plant canopy touches the top of the image frame.

The shooting distance, which varies according to plant height, is defined by Equation (4).

$$D = \frac{H}{2\tan(30°)} \tag{4}$$

where D is the optimal shooting distance from the camera to the base of the plant, and H is the height of the plant. The constant factor $1/2\tan(30°) \approx 0.866$ serves as a practical multiplier derived from trigonometric principles to estimate the necessary distance for field photography. As depicted in Figure 4.3, images are taken with a frame ratio of 2:1 from a height of 1.6 meters, ensuring that the top and bottom edges of the frame align with the top and bottom edges of the plant. This standardized approach provides a consistent basis for analyzing the visual features of urban green spaces.

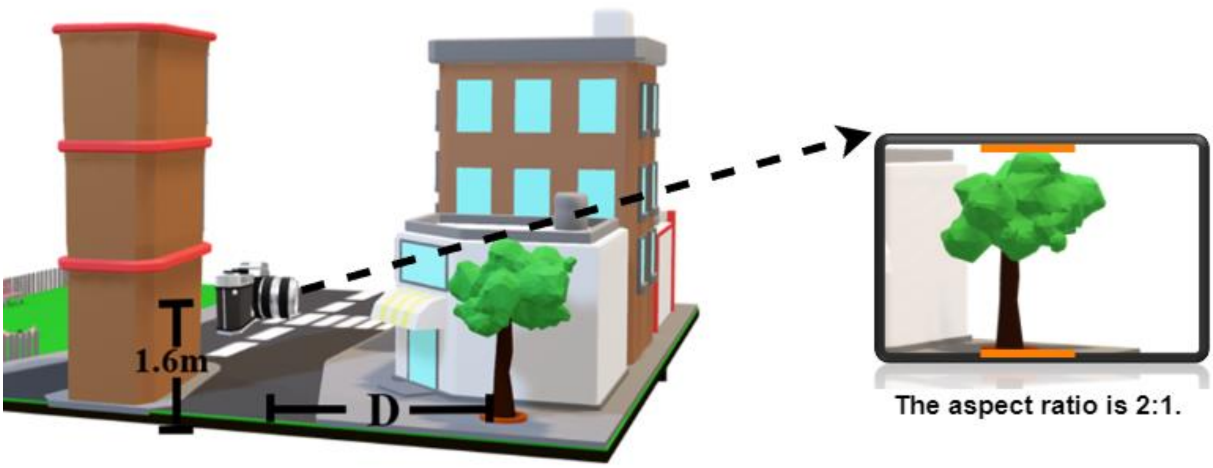


Figure 4.3 Standardized imaging method for plant analysis.

To ensure the usability of the view, Figure 4.4 compares the real street view images before and after preprocessing. The original street view images suffer from distortions and inappropriate perspectives, such as being taken from a distance or angles not representative of pedestrian viewpoints. After the 3D scene reconstruction and generation of standardized images, these issues are significantly improved.

(a) Original image: significant distortion, non-standard viewing angle, and distance.

(b) Preprocessed image: standardized plant observation angle and distance, improved image quality and consistency.

Figure 4.4 Comparison of street view images before and after preprocessing.

To evaluate the quality of the optimized images, the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are used (Hore & Ziou, 2010). PSNR and SSIM assessments reveal that the average PSNR of the preprocessed images increased by 5.2 dB and the average SSIM improved by 0.14. These improvements indicate that the preprocessing effectively reduces noise and enhances structural similarity, providing a high-quality data foundation for subsequent S3PVI calculations.

4.2.3 Plant visual feature evaluation based on S3PVI

In the multi-temporal urban green space vegetation visualization framework, the combination of image classification and semantic segmentation is a key step to accurately identify and separate the pixels of each plant species in standardized views. A deep learning-based image classification and semantic segmentation model is adopted, which is trained on a dataset containing 2,000 annotated images, covering 51 common plant species in urban environments, as shown in Table 4.2.

Table 4.2 IoU scores for common urban plant species in semantic segmentation

NUM	Botanical name	Common name	IOU (%)
1	Ternstroemia gymnanthera	Japanese Ternstroemia	90.74
2	Camptotheca acuminata	Cancer tree	91.70
3	Cupressus macrocarpa	Monterey cypress	89.04
4	Cinnamomum camphora	Camphor	81.38
5	Quercus acutissima	Quercus acutissima	70.97

NUM	Botanical name	Common name	IOU (%)
6	Ligustrum lucidum	Glossy privet	78.38
7	Acer	Maple	81.63
8	Gardenia jasminoides	Cape jasmine	89.85
9	Hibiscus makinoi	Makino's mallow	88.29
10	Pinus thunbergii	Japanese black pine	79.51
11	Cortaderia selloana	Pampas grass	86.74
12	Paliurus ramosissimus	Thorny wingnic	79.07
13	Camellia japonica	Japanese camellia	77.34
14	Acacia baileyana	Cootamundra wattle	88.23
15	Lithocarpus edulis	Japanese stone oak	80.93
16	Castanopsis sieboldii	Itajii	73.29
17	Torreya nucifera	Japanese torreya	77.38
18	Quercus myrsinifolia	Japanese white oak	87.37
19	Ginkgo biloba	Ginkgo	74.79
20	Bassia scoparia	Kochia	81.71
21	Rhaphiolepis indica var. umbellata	Rhaphiolepis umbellata	81.64
22	Rosa spp.	Rosa	86.41
23	Cordyline spp.	Cordyline	79.72
24	Quercus glauca	Ring-cupped oak	87.55
25	Ceratonia siliqua	Arakashi	73.81
26	Jacaranda mimosifolia	Blue Jacaranda	83.17
27	Washingtonia filifera	California palm	80.21
28	Erythrina bidwillii	Coral tree	86.92
29	Paeonia lactiflora	Chinese peony	89.65
30	Styphnolobium japonicum var. pendulum	Japanese pagoda	84.75
31	Lavandula angustifolia	English lavender	73.21
32	Pelargonium crispum	lemon geranium	92.50
33	Salvia rosmarinus	Rosemary	90.48
34	Litsea japonica	Hamabiwa	82.25
35	Chamaecyparis pisifera 'Filifera'	Sawara cypress	69.48
36	Rhododendron spp.	Azalea	79.39
37	Eurya emarginata	Eurya emarginata	87.23
38	Juniperus rigida	Temple juniper	75.00
39	Cycas revoluta	Sago palm	83.68
40	Photinia・fraseri	Christmas berry	85.22
41	Muhlenbergia capillaris	Muhly grass	90.78
42	Magnolia denudata	Lily tree	76.77
43	Prunus serrulata	Cherry blossoms	69.61
44	Picea abies	Norway spruce	81.13
45	Helianthus annuus	Common sunflower	75.49
46	Hedera canariensis	Canary ivy	87.57
47	Brassica oleracea var. acephala	Acephala group	91.82

NUM	Botanical name	Common name	IOU (%)
48	Forsythia suspensa	Weeping forsythia	74.69
49	Osmanthus fragrans	Sweet osmanthus	87.02
50	Pinus pinea	Stone pine	86.05
51	Pittosporum tobira	Japanese cheese wood	69.10

1) Dataset development and characteristics:

The dataset was meticulously curated, drawing inspiration from the cityscapes dataset (Cordts et al., 2016). Special attention was given to ensuring a diverse representation of plant types and environmental conditions, with a particular focus on the ornamental status of plants. Each image underwent a rigorous annotation process, involving at least two individuals to provide pixel-level labels. The dataset was strategically divided into training, validation, and test subsets at a ratio of 70%, 15%, and 15%, respectively (Gomes & Zheng, 2020).

2) Model architecture and selection:

The principle of image classification is distinguishing between object classes by processing their distinct features in an image. Deep neural networks, such as convolutional neural networks and visual transformers, have shown remarkable performance, especially when trained on large, well-labeled datasets (Kolesnikov et al., 2020). In this analysis, an image classification model was first applied to streamline and enrich the subsequent semantic segmentation. The EfficientNet architecture (Tan & Le, 2020) was used for tree species classification, with the EfficientNet-b4 variant achieving a commendable accuracy of 97.9% on this dataset.

For semantic segmentation, which labels each pixel in an image, several models were tested. PSPNet (Zhao et al., 2017), DANet (Xue et al., 2019), and ISANet (L. Huang et al., 2019) emerged as top performers. The models were evaluated using mean square error (MSE), with DANet scoring the lowest MSE of 148.74, outperforming ISANet (235.66) and PSPNet (337.09). Considering all evaluations, EfficientNet-b4 was selected for image classification, while DANet was chosen for the segmentation task.

3) Training process and data augmentation:

The model was trained using the Adam optimizer with an initial learning rate of 0.0001, a batch size of 8, and a total of 100 epochs (Kingma & Ba, 2017). Various data augmentation techniques, including random flipping, scaling, and cropping, were implemented to enhance the model's robustness and generalization ability (Shorten & Khoshgoftaar, 2019).

4) Model performance and analysis:

The final model achieved a mean Intersection over Union (mIoU) of 82.17% on the test set, demonstrating its effectiveness in distinguishing different plant species. Table 4.2 presents a detailed breakdown of the IoU scores for each of the 51 plant species. The scores range from 69.10% to 92.50%, highlighting varying performance across different species.

Lemon geranium achieved the highest IoU score of 92.50%, indicating exceptionally accurate segmentation. Other high-performing species include cancer tree (91.70%), acephala group (91.82%), and Japanese ternstroemia (90.74%).

At the lower end, Japanese cheese wood had the lowest IoU of 69.10%, followed closely by cherry blossoms at 69.61% and Sawara cypress at 69.48%. These lower scores suggest challenges in accurate segmentation for these species.

The case of cherry blossoms (69.61% IoU) warrants deeper analysis. Despite focusing on images of blooming cherry trees, the relatively low performance could be attributed to several factors.

Variability in flower density: Cherry blossoms can vary significantly in their blooming density, challenging the model's ability to consistently segment the flowers.

Viewing angles: Images taken from various angles (below, eye level, or above) could affect the model's ability to consistently identify and segment the blossoms.

Cultivar diversity: Numerous cherry blossoms cultivars with subtle differences in flower shape, size, and color may not be adequately represented in the training data, leading to lower segmentation accuracy for some varieties.

Other species with lower IoU scores, such as Japanese white oak (74.79%) and lily tree (76.77%), may face similar challenges related to variability in appearance or similarity to other species. For instance, Japanese white oak might be confused with other oak species, while lily tree could be challenging due to its seasonal changes in appearance.

These results highlight the complexity of urban vegetation classification and segmentation, where factors such as species variability, environmental conditions, and image capture techniques can significantly influence plant appearance and model performance.

To address these challenges, future work could focus on expanding the dataset for challenging species, ensuring a wider representation of varieties and growth conditions. Incorporating multi-temporal data to capture seasonal variations, especially for species with significant appearance changes, would be beneficial. Developing species-specific sub-models or fine-tuning techniques for more accurate segmentation of challenging species could significantly improve performance. Exploring advanced data augmentation techniques to better simulate real-world variability in plant appearances would enhance the model's robustness. Additionally, investigating the use of ensemble methods or multi-scale approaches could improve segmentation accuracy across diverse species. These improvements would collectively contribute to enhancing the model's performance across all plant species in urban environments, further solidifying its utility in urban green space analysis.

While the model's overall performance is strong, with a high mean IoU, the analysis reveals areas for potential improvement, particularly in handling species with high variability or complex morphological characteristics. The trained model is applied to the standardized views generated from the reconstructed 3D scene to generate segmentation maps.

Figure 4.5 compares original photographs of various plant species with their corresponding segmented images. Each row showcases different species, such as cherry blossoms, false cypress, maple, and kochia. The segmented images use distinct colors to represent each plant species: pink for cherry blossoms, light blue for azalea, yellow for false cypress, green for rose, brown for lemon geranium, light gray for pampas grass, reddish-brown for maple, purple for kochia, white for California palm, beige for ginkgo, and dark gray for Japanese black pine.

The segmentation accurately identifies and separates the plant species from the background, highlighted in black, demonstrating the effectiveness of the segmentation process in differentiating between various types of vegetation in diverse settings. For example, the image clearly shows the segmentation of cherry blossoms in pink, a tall maple tree in reddish-brown, and bright kochia plants in purple.

All plants in the dataset can be segmented in the segmentation model, as evidenced by the variety of species represented in the image, from trees like false cypress and ginkgo to shrubs like rose and grasses like pampas grass. The segmentation image can be used to calculate the S3PVI value and subsequently analyze the changes in the visual visualization features of plants, which is particularly useful given the diverse range of plant types depicted in the photographs.

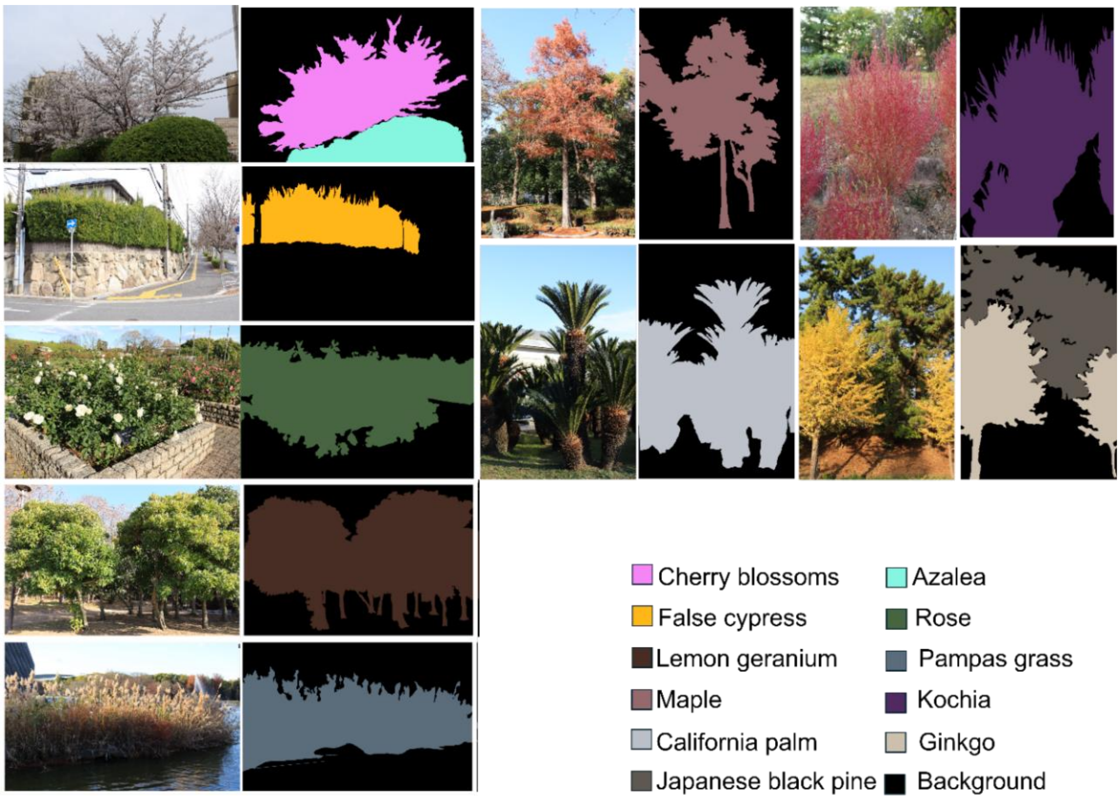


Figure 4.5 Examples of original and segmented images.

4.2.4 S3PVI calculation for plant visual feature quantification

To ensure consistent and reliable data collection for S3PVI evaluation, standardized field photography procedures have been developed, drawing from best practices in urban vegetation assessment and visual quality analysis (Osaka Prefectural Government, 2020).

After semantic segmentation, the S3PVI of each plant species in the segmented image is calculated using the Equation (5):

$$S3PVI=\frac{\sum_{i=1}^{n} Area_{v_i}}{\sum_{i=1}^{n} Area_{t_i}} \times 100\ (\%) \tag{5}$$

where n is the total number of photos taken in the test area, $Area_{v_i}$ is the total number of pixels of the target plant in image i took along the horizontal direction, and $Area_{t_i}$ is the total number of pixels in image i. This ratio reflects the percentage of pixels attributed to the plant relative to the entire image, providing a S3PVI value from 0% to 100%. The average of these values for all images quantifies the plant's visibility in the area.

The S3PVI is inspired by the GVI, which quantifies vegetation visibility in street-level images. While the GVI focuses on overall greenery visibility, the S3PVI extends this concept by quantifying the visibility of individual plant species across multiple seasons. This species-level, multi-temporal analysis provides a more detailed understanding of the visual characteristics and dynamics of urban green spaces. The average S3PVI values across all images quantify the visibility of each plant species in the study area. By evaluating multiple species in each image, the S3PVI enables a detailed assessment of their respective contributions to the aesthetics of urban green spaces.

4.3 Experiments and results

4.3.1 Visual feature analysis of real street vegetation

To validate the effectiveness of the multi-temporal urban green space vegetation visualization framework, Sanshikisaido in Suita City, Osaka Prefecture, Japan, was selected as a case study. Sanshikisaido is situated in the Kita-senri residential area of Suita City, extending approximately 500 meters. It is renowned for its seasonal foliage displays, particularly its autumn colors. Figure 4.6 provides a detailed overview of Sanshikisaido's vegetation composition in Suita City, Osaka Prefecture, Japan.

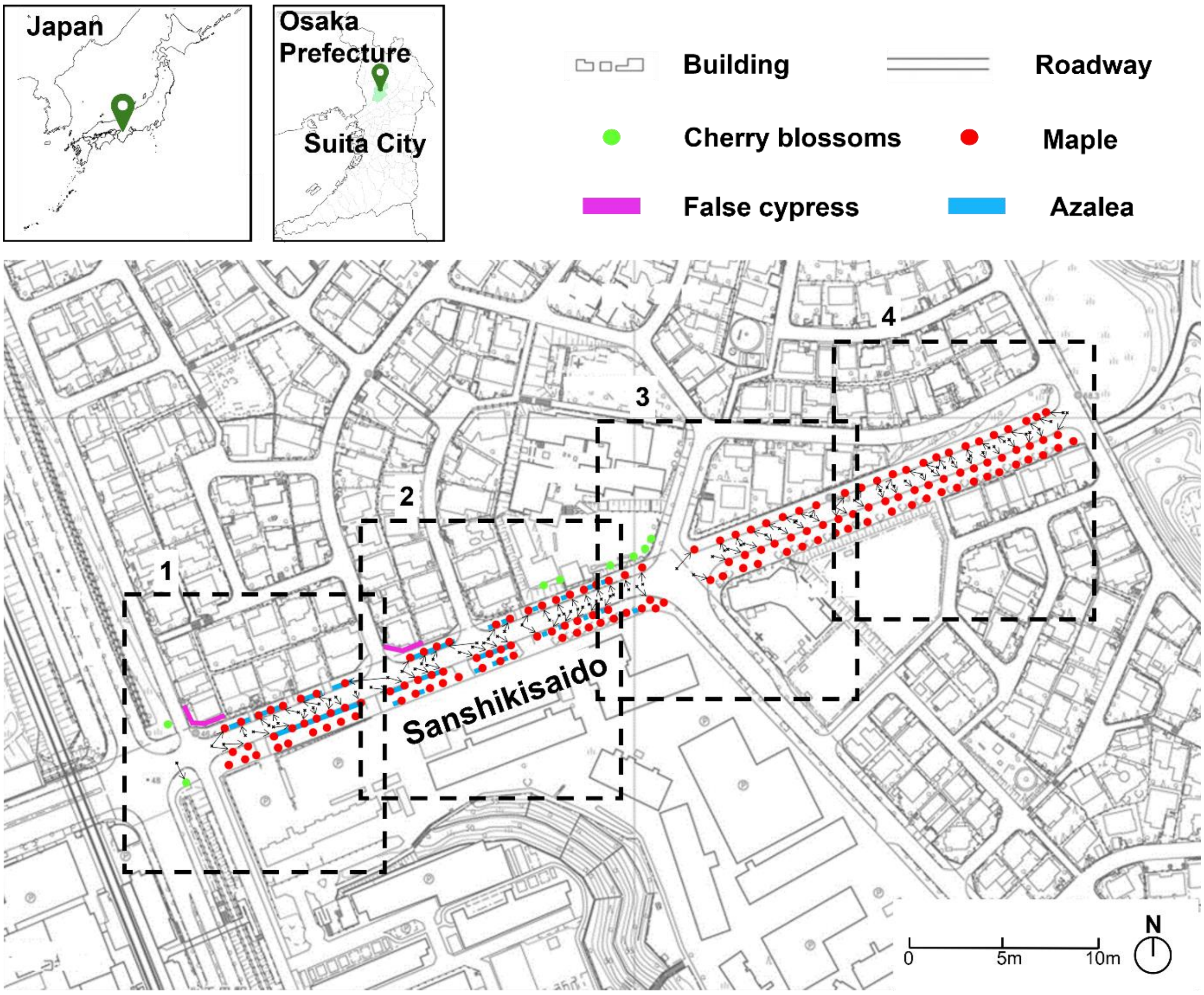


Figure 4.6 Vegetation distribution along Sanshikisaido, Suita City.

The map focuses on the Sanshikisaido road and its surrounding urban area, with a legend explaining symbols for buildings, roadways, and various plant species. The central feature is Sanshikisaido itself, lined predominantly with maple trees and interspersed with cherry blossoms. The urban layout surrounding the road is clearly defined, showing buildings and streets. The area is divided into four numbered sections by dashed lines.

Figure 4.7 shows zone 1 lined with various plants, predominantly maple trees (represented by red dots). There are also several azalea bushes (blue lines) interspersed among the maples. A cluster of false cypress trees (magenta line) is visible on one side of the street, and two cherry blossoms trees (green dots) are located at opposite ends of the depicted area. This close-up includes additional information about camera placements. Sixteen camera locations are marked with small 'x' symbols, numbered from 1 to 16. Each camera location is accompanied by an arrow indicating the camera's orientation or direction of view. The legend in the bottom right corner of this detailed view confirms the symbols used for different vegetation types and adds explanations for the camera-related markings.

Figure 4.7 Detailed vegetation map and camera placement in zone 1 of Sanshikisaido.

Figure 4.8 showcases zone 2 of Sanshikisaido, extending the vegetation mapping from the previous figure. The street's vegetation remains predominantly maple trees, with interspersed azaleas. Notable additions include a small cluster of false cypress trees on the left and several cherry blossoms in the upper right corner. Camera placements continue from the previous zone, numbered 16 to 40, each marked with its orientation.

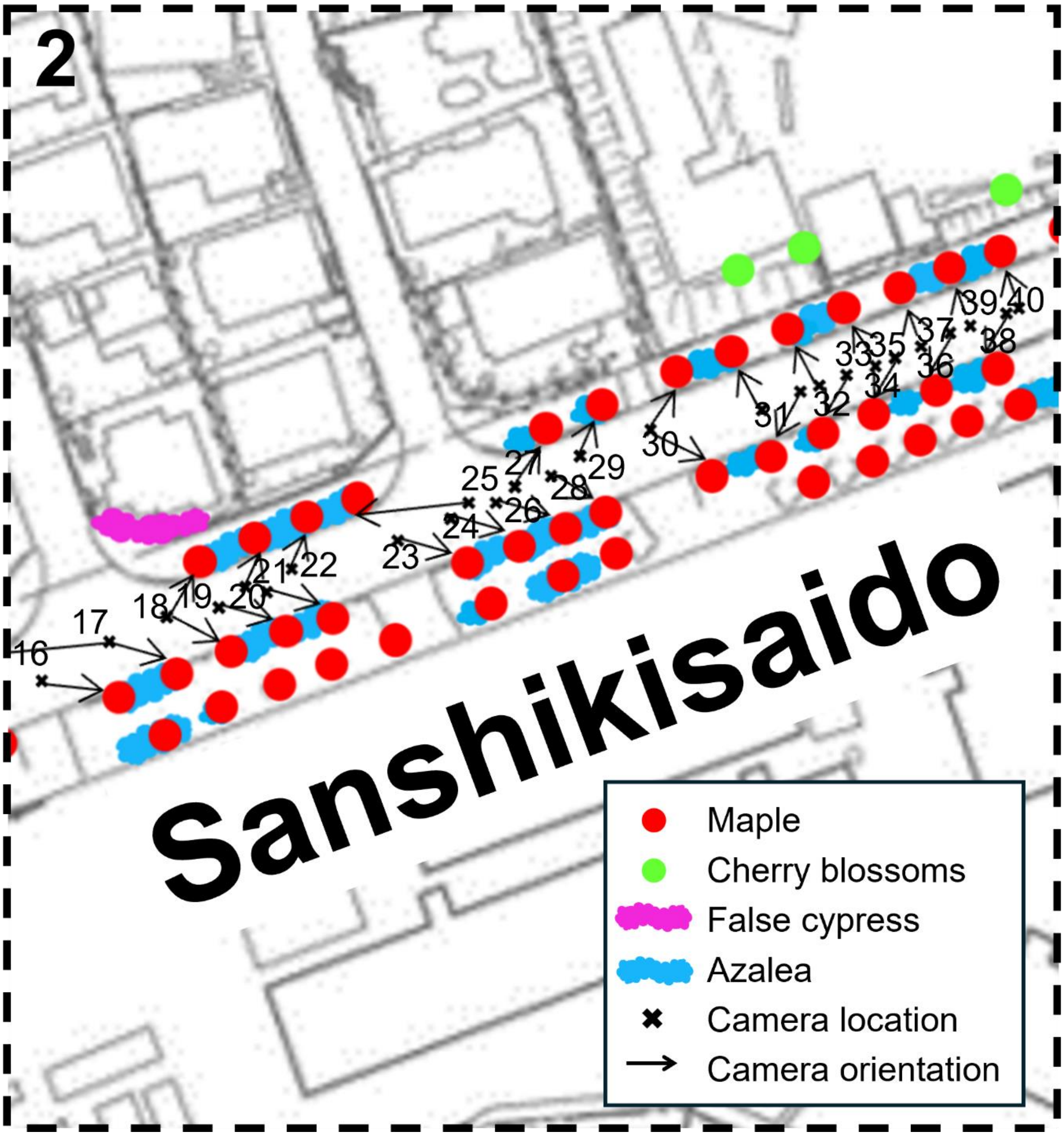


Figure 4.8 Detailed vegetation map and camera placement in zone 2 of Sanshikisaido.

Figure 4.9 illustrates zone 3 of Sanshikisaido. This section shows a curve in the street, with maple trees remaining the dominant vegetation. A notable cluster of cherry blossoms appears on the curved portion of the road. Camera locations are numbered from 39 to 58, following the street's contour.

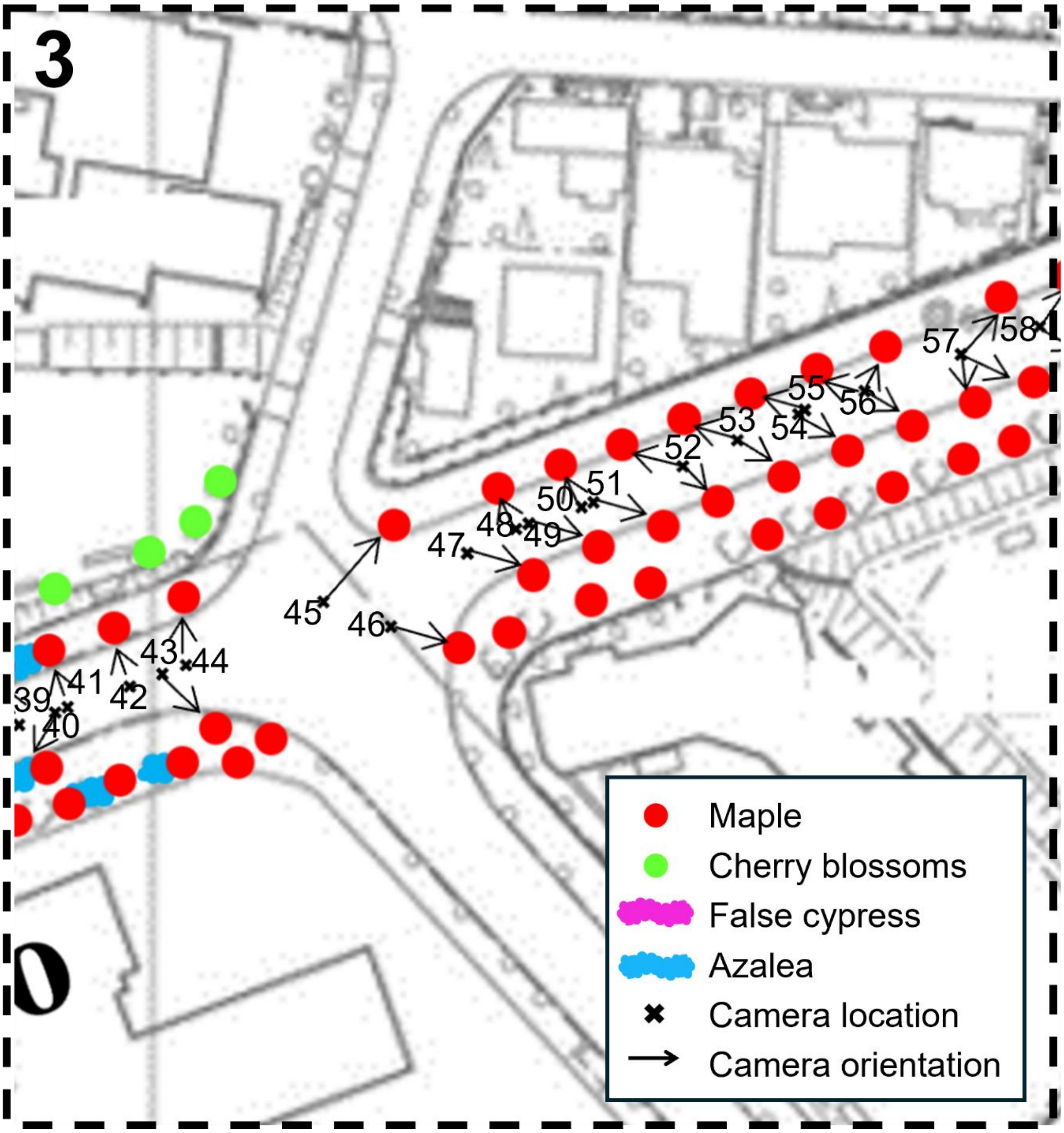


Figure 4.9 Detailed vegetation map and camera placement in zone 3 of Sanshikisaido.

Figure 4.10 depicts zone 4 of Sanshikisaido, the final section of the street study. This zone shows a gently curving road lined exclusively with maple trees, in contrast to the more diverse vegetation in previous zones. Camera positions are numbered from 58 to 80, following the street's curve.

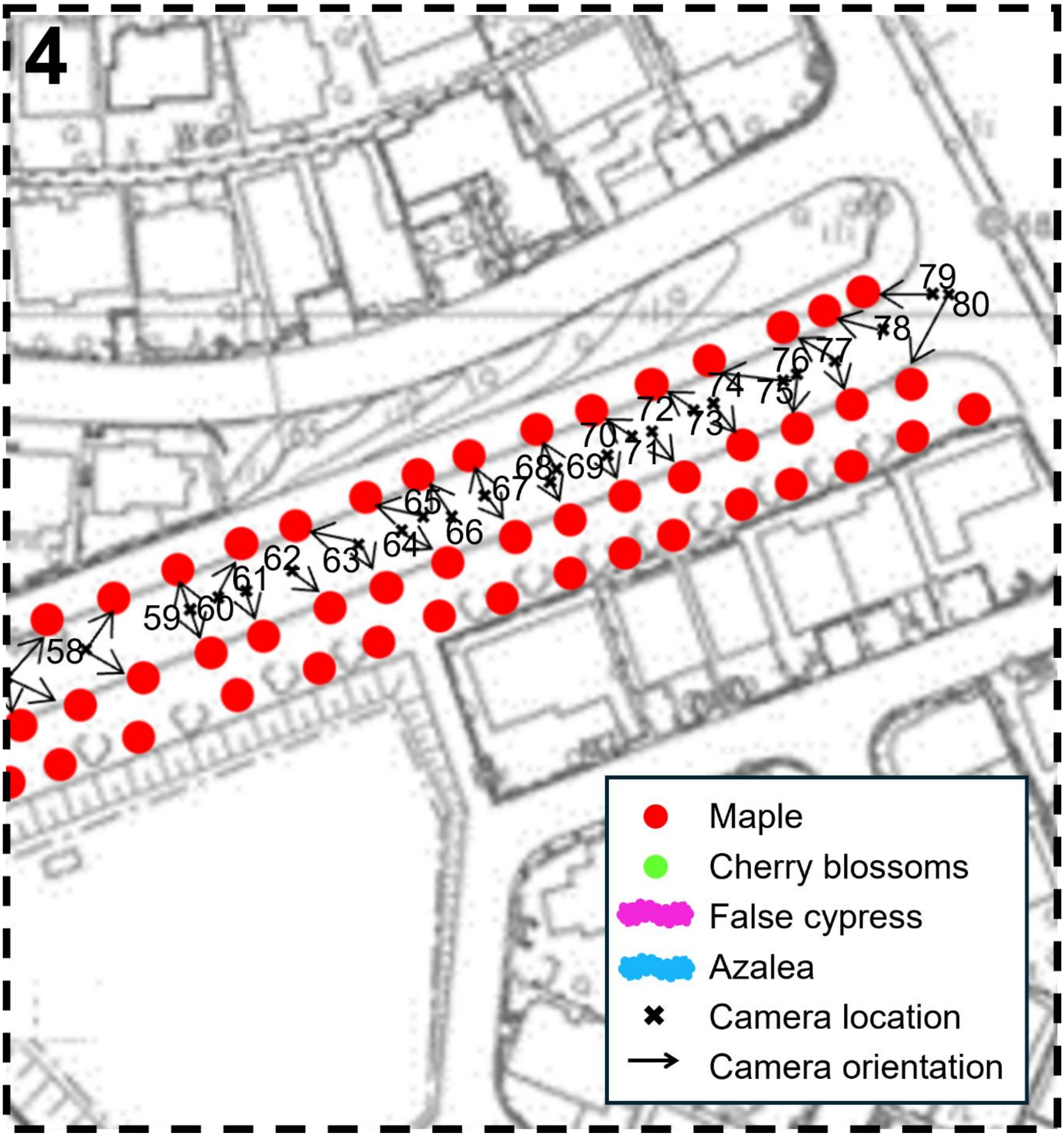


Figure 4.10 Detailed vegetation map and camera placement in zone 4 of Sanshikisaido.

The 80 digital camera icons represent locations where standardized view generation was implemented. Each camera location adheres to the height and distance specifications outlined in Figure 4.3, ensuring consistent and comparable photographic data collection across the entire Sanshikisaido street. This standardized approach is crucial for accurate temporal analysis and comparison of urban vegetation changes.

From these 80 locations, four representative sites (Locations 1, 2, 25, and 45) were selected to showcase the diverse vegetation types and their temporal dynamics along Sanshikisaido.

To access and analyze these historical street view data, the Google Maps built-in time travel feature was utilized. This tool allows users to view historical imagery by selecting the street view date in the upper-left corner of the interface, which opens a time slider. By manipulating this slider, images from different years can be accessed, enabling a comprehensive temporal analysis of the streetscape. This method facilitated systematic access to and comparison of street View images from different periods, allowing for an in-depth spatiotemporal analysis.

Figure 4.11 shows Google Street View images collected at Location 1 for 10 different periods between 2010 and 2022. This comprehensive dataset represents all available Street View data for the region, ensuring a thorough temporal analysis. The image collection follows the standardized view generation protocol described in Section 4.2.2, keeping the camera angle, height, and distance consistent across all time periods. All photo data was input into a multi-temporal evaluation framework to obtain the corresponding S3PVI values. The lower half of Figure 4.11 shows the identified plants and the corresponding segmented images for each photo. The cherry blossoms are highlighted in pink, while the background is black. This visual representation makes it easy to track the coverage of cherry blossoms over time.

Figures 4.12, 4.13, and 4.14 present similar temporal analyses for Locations 2, 25, and 45 respectively.

Location 2 (Figure 4.12) focuses on false cypress trees. The segmented images show these trees in yellow, with consistent presence across all time periods. A notable maple presence is also detected in the 2018.11 image.

Location 25 (Figure 4.13) emphasizes azalea growth. The segmented images highlight azaleas in light blue, showing variations in their presence and size over the years.

Location 45 (Figure 4.14) showcases the evolution of maple trees. The segmented images highlight maple foliage in reddish-brown, particularly visible in the 2018.11 image, capturing autumn colors.

These additional locations demonstrate the versatility of the standardized view generation method in capturing diverse urban vegetation types and their temporal changes. The consistent imaging and segmentation approach across all locations enables comparative analysis of different plant species' growth patterns and seasonal variations.

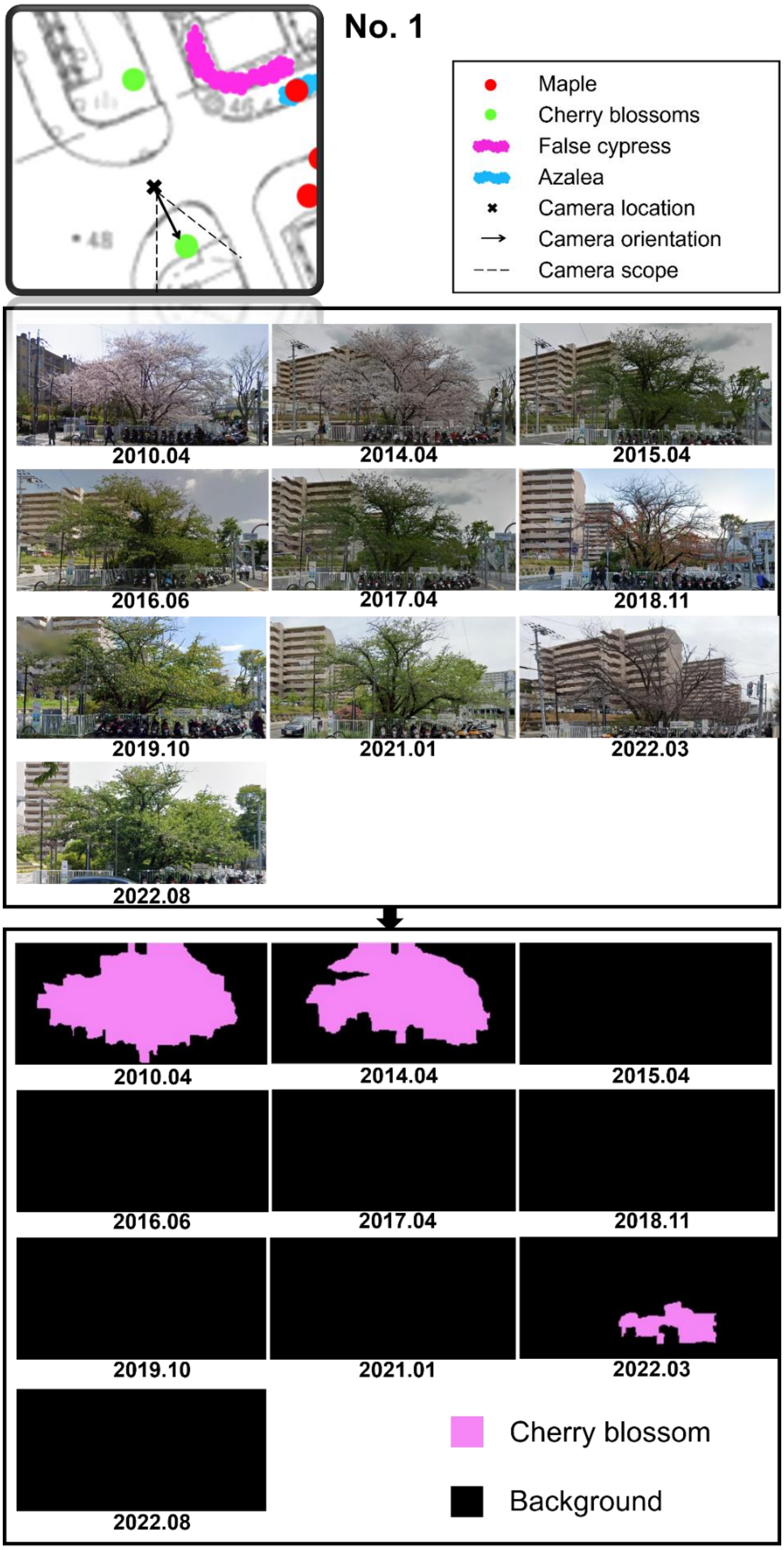


Figure 4.11 Temporal analysis of vegetation presence at location 1 (2010-2022)

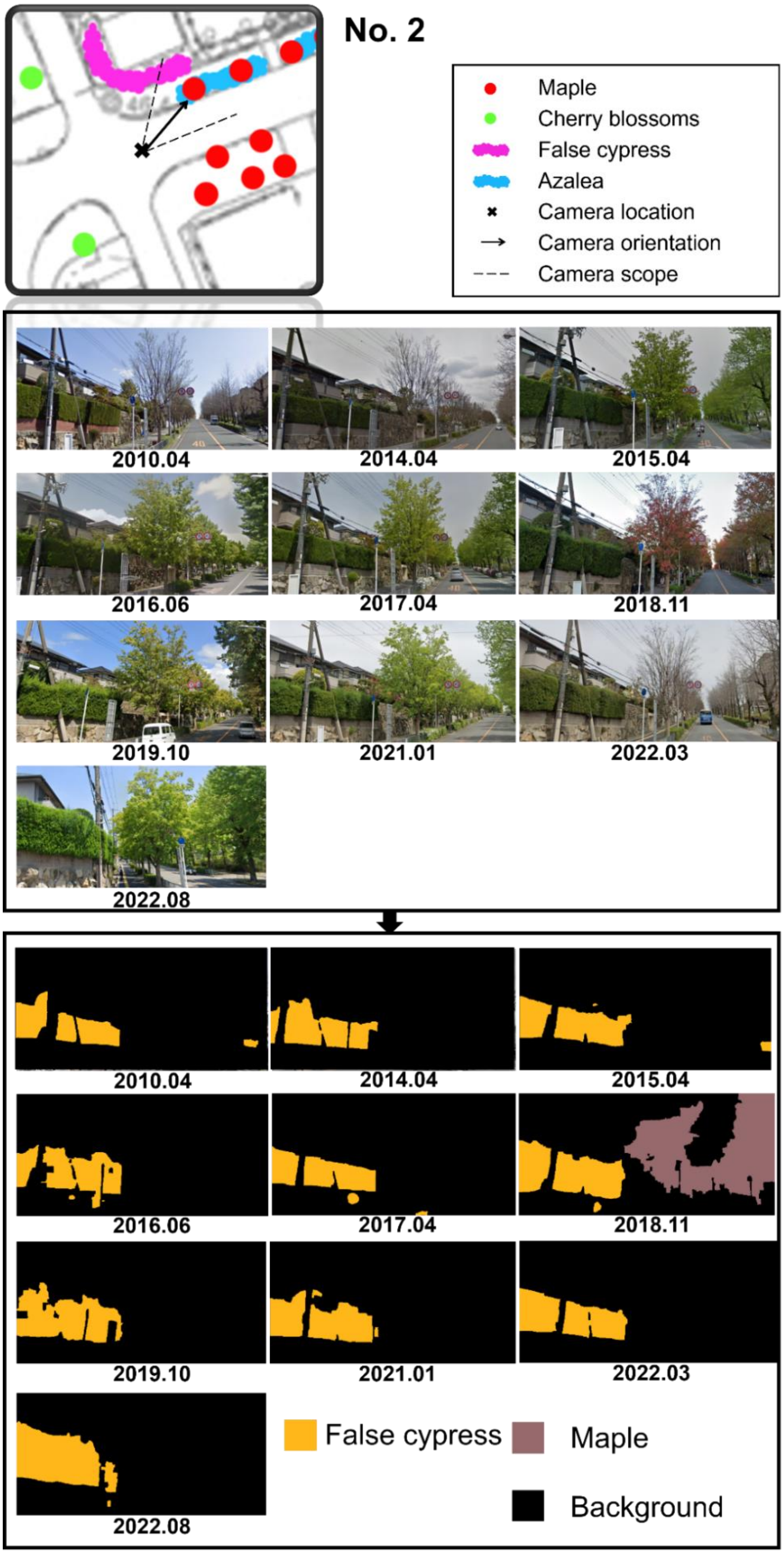


Figure 4.12 Temporal analysis of vegetation presence at location 2 (2010-2022)

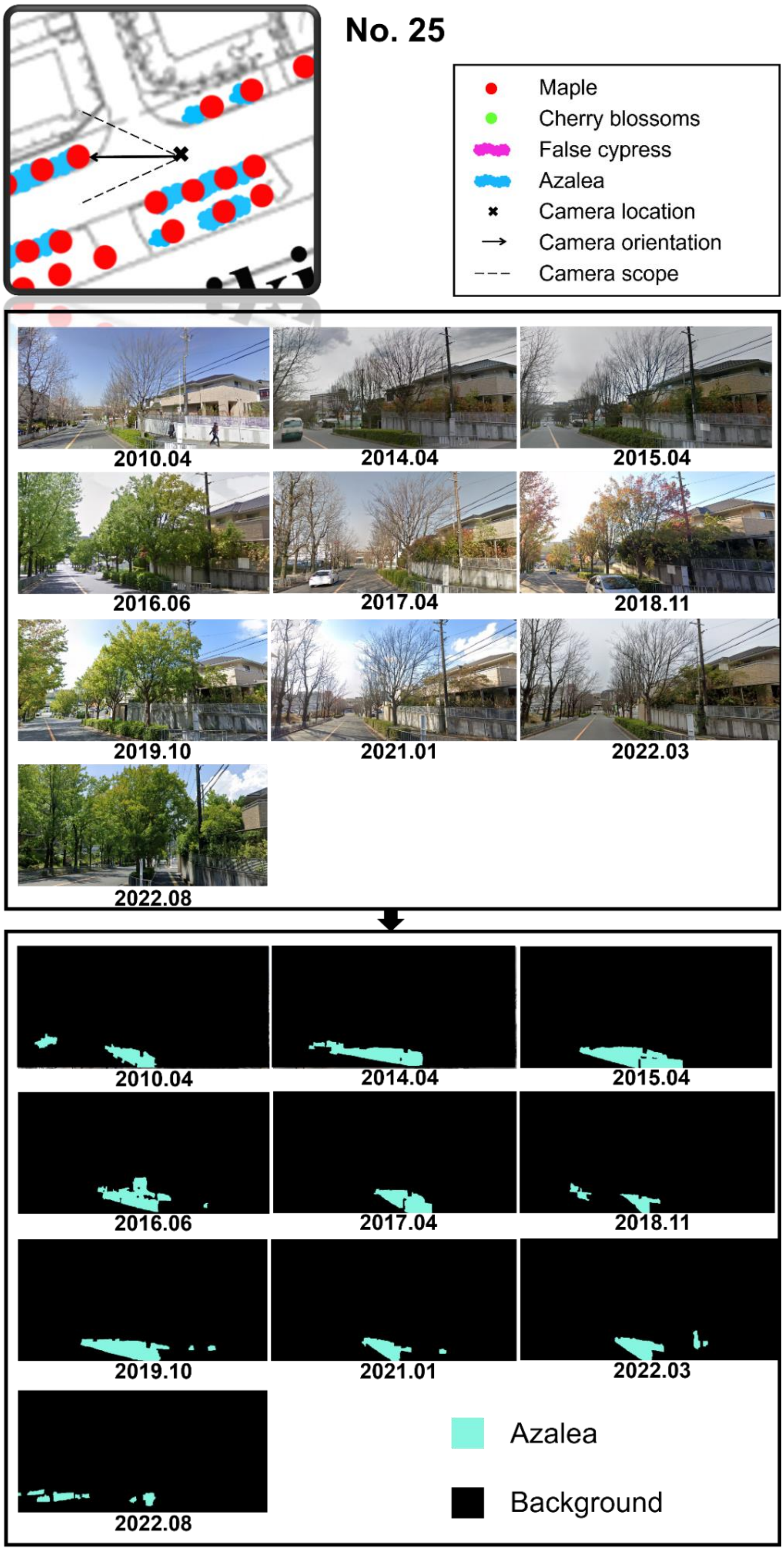


Figure 4.13 Temporal analysis of vegetation presence at location 25 (2010-2022)

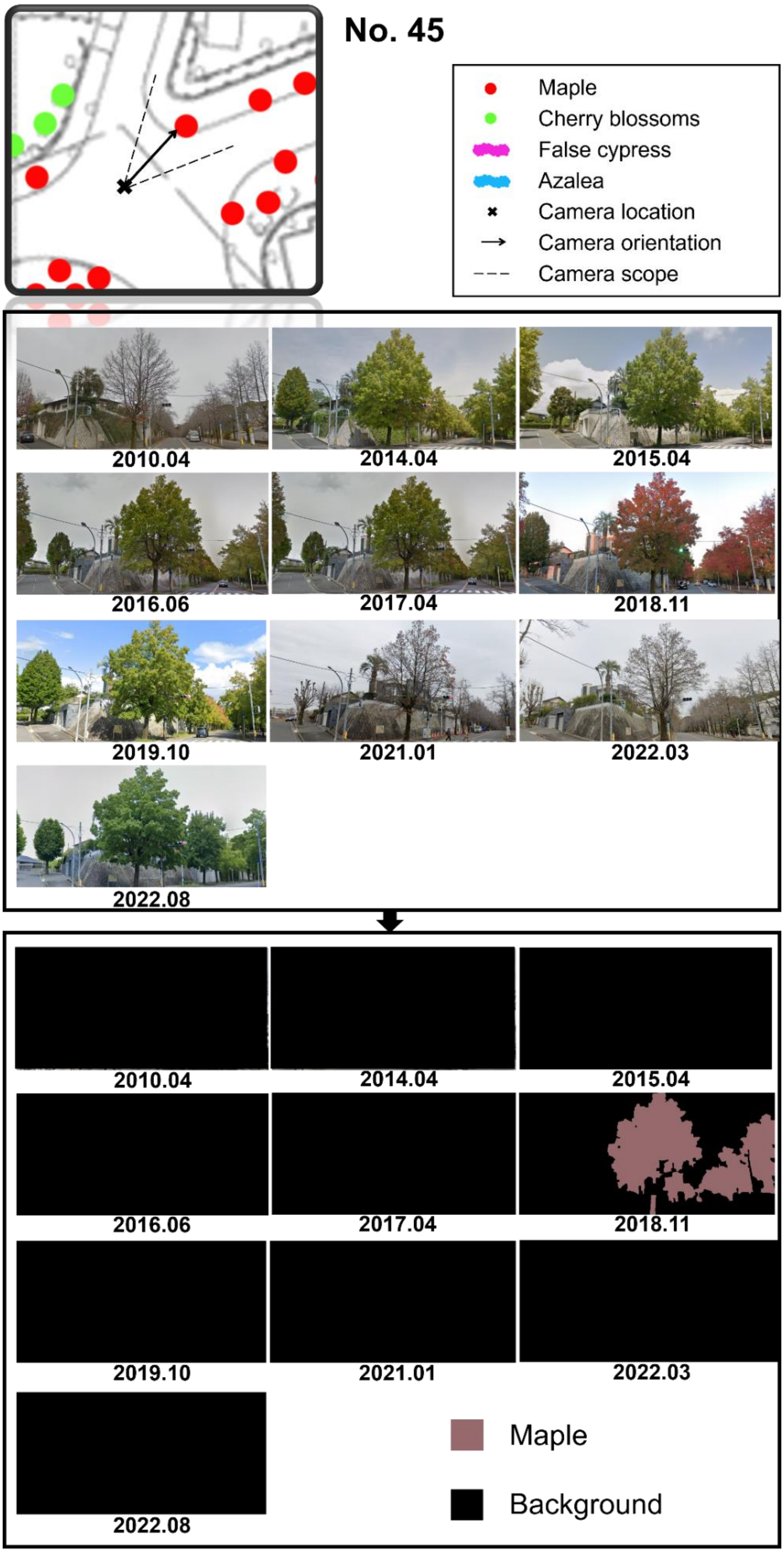


Figure 4.14 Temporal analysis of vegetation presence at location 45 (2010-2022)

In the street view data of Sanshikisaido, the S3PVI values of four major plant species were detected, including cherry blossoms, maple, false cypress, and azalea. It's important to note that the current focus is on the aesthetic viewing periods of plants, particularly when they have special ornamental value. As a result, the dataset of 2000 plant images and subsequent S3PVI calculations are primarily based on times when plants are in their most visually appealing states. For instance, in the case of Sanshikisaido, which is famous for its seasonal foliage display, the current S3PVI values are calculated during these optimal aesthetic periods.

This approach enables the quantification and visualization of the seasonal peak aesthetic value of urban green spaces, with a particular emphasis on iconic or culturally significant vegetation periods. Figure 4.15 presents a comprehensive 3D visualization of the S3PVI values for four key plant species in Sanshikisaido from 2010 to 2022. The graph's axes represent time (x-axis), plant species (y-axis), and S3PVI values in percentage (z-axis). The four species tracked are cherry blossoms (red line), maple (orange line), false cypress (green line), and azalea (blue line).

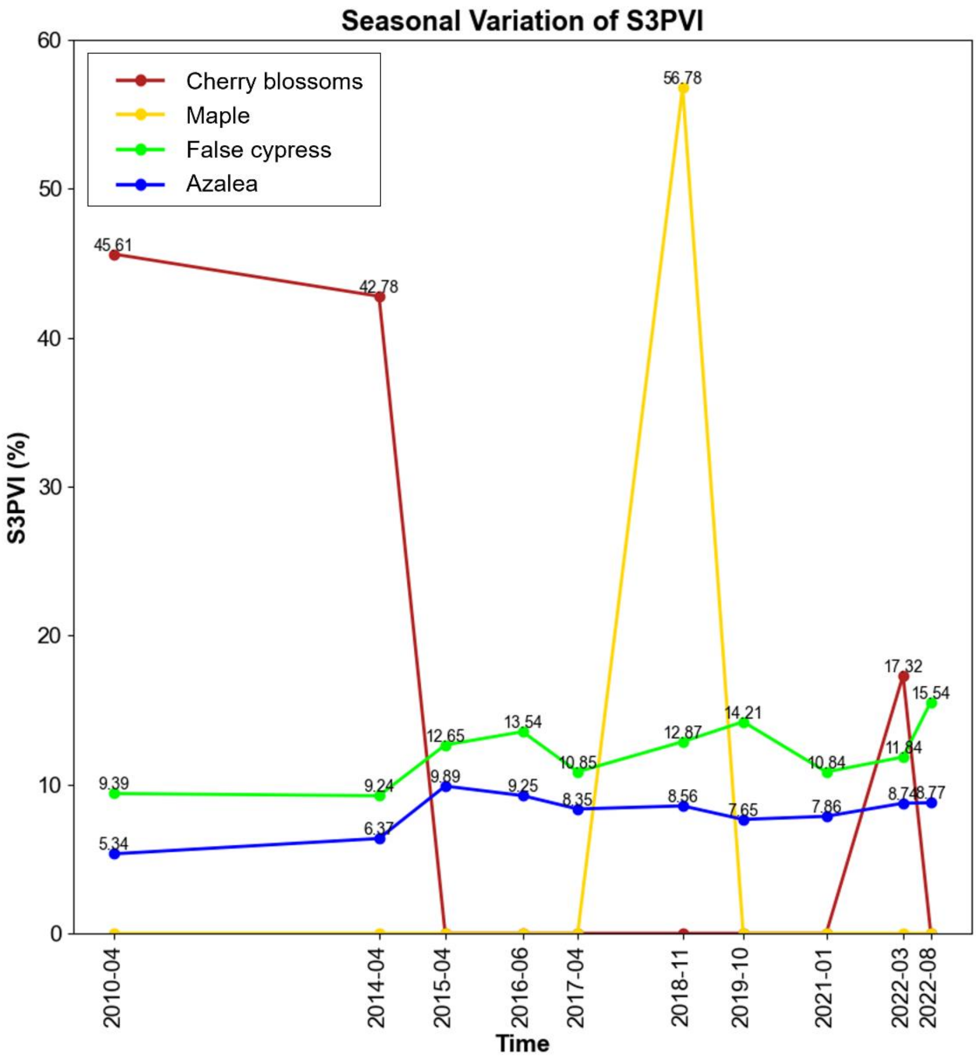


Figure 4.15 Seasonal change curves of S3PVI values for major plant species

Cherry blossoms show distinct peaks in April 2010 (45.61%) and April 2014 (42.78%). However, from 2015 onward, the S3PVI values for cherry blossoms are notably lower. It's important to note that these lower values do not indicate the absence of cherry trees, but rather suggest that they were not in their peak blooming state during the observed periods.

Maple exhibits the highest variability among the four species. Its S3PVI value reaches a maximum of 56.78% in November 2018, likely corresponding to its autumn foliage display. The maple's S3PVI values show clear seasonal patterns, with higher values typically occurring in autumn months and lower values at other times of the year.

False cypress demonstrates relatively consistent S3PVI values throughout the observed period, generally ranging between 10% and 15%. This stability reflects the evergreen nature of the species, maintaining its visual presence across seasons.

Azalea's S3PVI values fluctuate moderately, generally ranging from about 5% to 10%.

Figure 4.15 effectively captures the seasonal dynamics of these four species' visual impact in Sanshikisaido's urban landscape. It highlights the periods when each species contributes most significantly to the aesthetic value of the area, while also revealing patterns in their visual prominence over the years.

As the visual dominance transitions from cherry blossoms in early spring to false cypress and azalea in late spring and summer, and then to maples in autumn, we see a landscape that constantly evolves, offering diverse visual experiences throughout the year. This pattern not only highlights the seasonal strengths of each species but also points to the system's methodology in assessing visual contributions, which appears to be particularly attuned to capturing peak aesthetic moments of each plant type.

The S3PVI tool provides data for landscape planners and designers, particularly in evaluating current conditions, especially in brownfield projects. It offers objective information on the visibility of different plant species throughout their ornamental seasons, which can be used as a baseline for further analysis and decision-making. Planners and designers can use these insights in conjunction with their professional knowledge, on-site investigations, and consideration of local ecological and social factors to develop urban green space strategies.

4.3.2 Multi-temporal visualization evaluation of virtual park vegetation design schemes

To validate the practical application of the multi-temporal urban green space vegetation visualization framework in urban planning and design, this study conducted a simulation experiment using a virtual park design.

Figure 4.16 presents a detailed plan view of this virtual park, showcasing a thoughtfully designed 7,000 square meter urban green space. This compact yet diverse layout incorporates a range of landscape elements, including a water feature, various structures, a winding pedestrian pathway, and a rich variety of vegetation. The design strategically positions trees, shrubs, and herbs to create distinct zones and visual interest, while also incorporating landscape stones and different ground covers. This virtual model serves as an ideal testbed for simulating the park's visual evolution over time, demonstrating how the multi-temporal vegetation visualization framework can be applied to urban planning and design processes, helping stakeholders envision and optimize green spaces as they develop and mature.

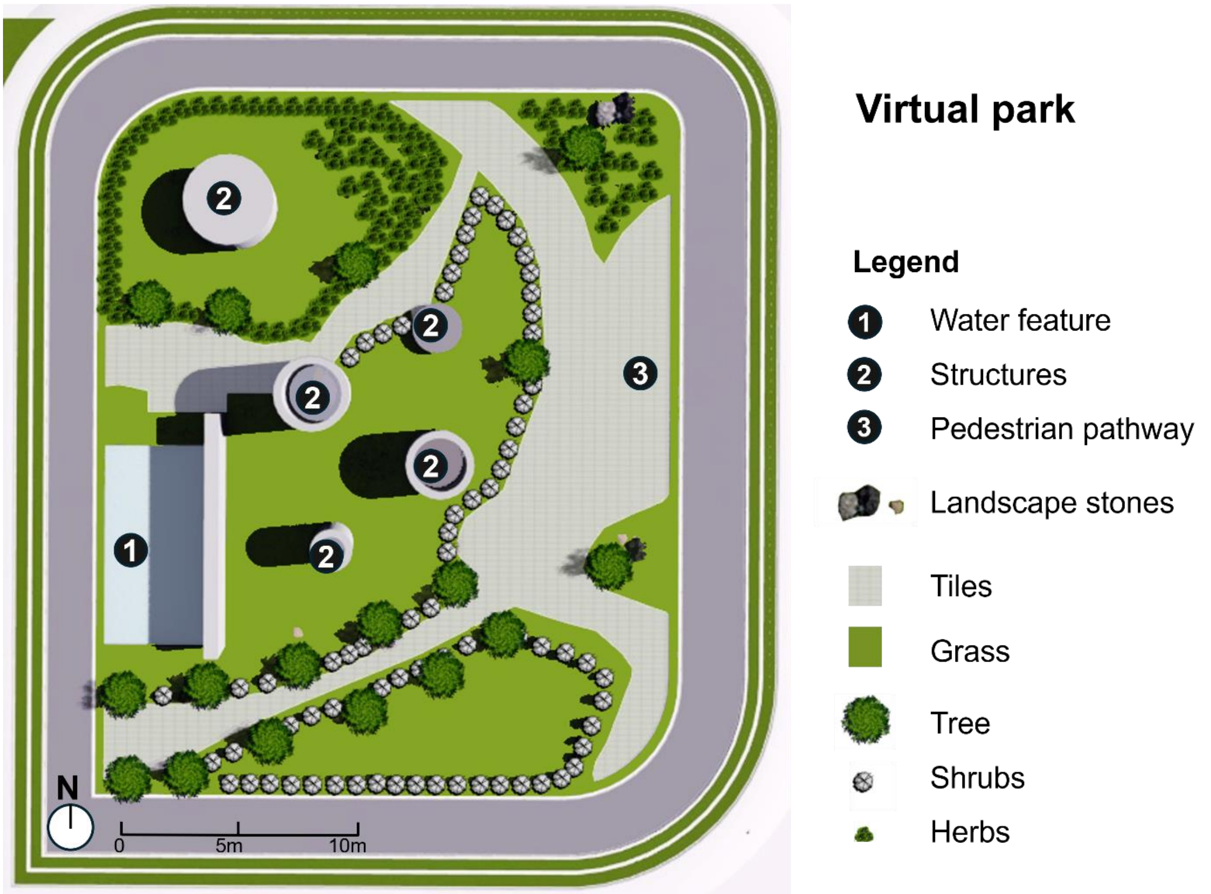


Figure 4.16 Virtual park design layout for multi-temporal vegetation visualization.

Figure 4.17 provides a comprehensive layout of the 14 specific camera locations in the virtual park design. The figure illustrates the strategic placement of these camera locations, showing their distances from the target plants, and the corresponding vegetation views seen from each point.

The left side of Figure 4.17 details the positions of the 14 viewpoints within the park, each marked with a number and an arrow indicating the direction of the view. These points are carefully distributed throughout the park to capture a diverse range of visual perspectives. The distances between each viewpoint and the target vegetation are indicated, providing precise measurements that are essential for accurate visual analysis.

On the right side of the Figure 17, each viewpoint is associated with a specific plant view, depicting the vegetation visible from that point. These visual representations include detailed images of the plants, highlighting their species, size, and arrangement.

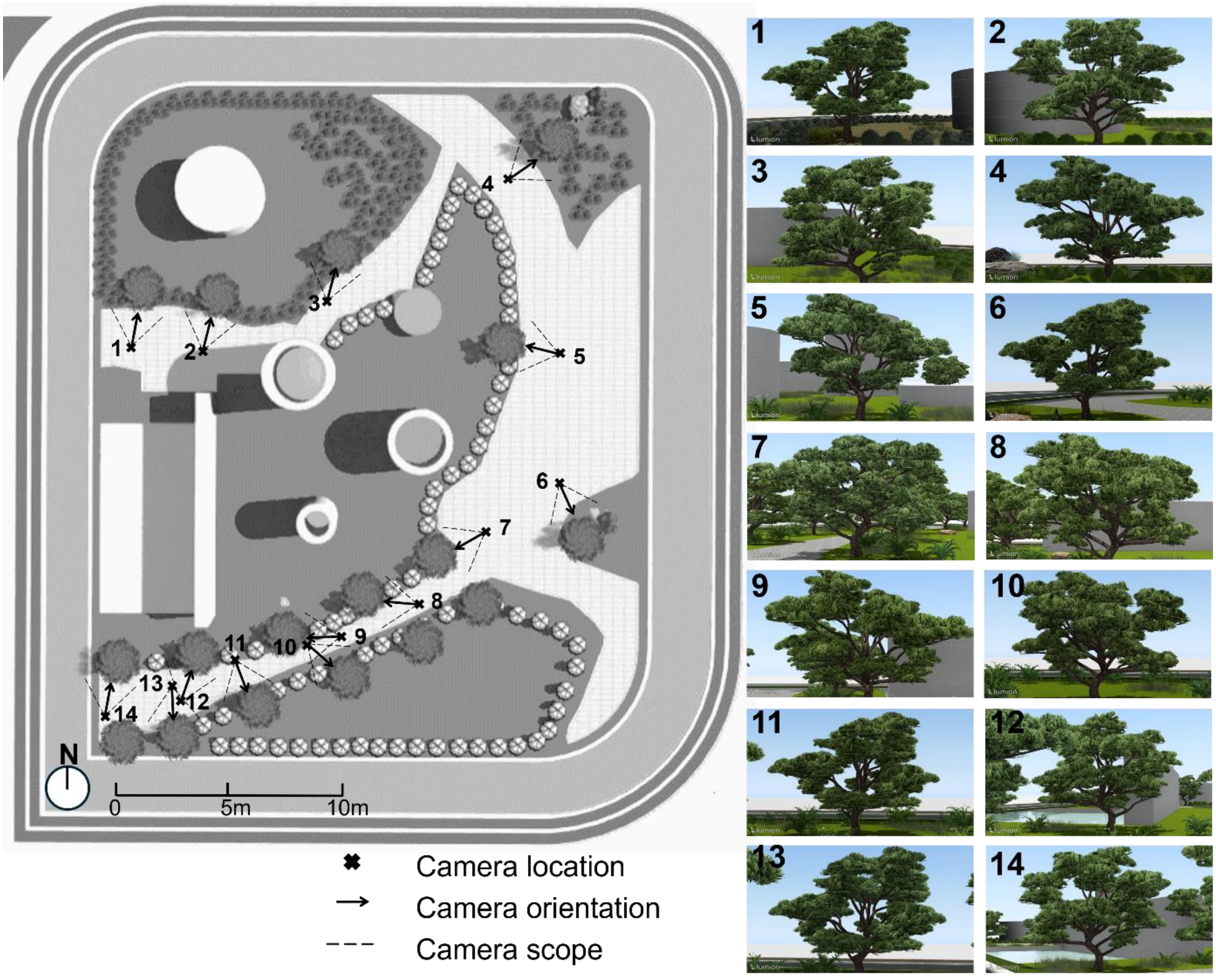


Figure 4.17 Location map of the virtual camera setup for the virtual park.

Figure 4.18 illustrates seven different planting design schemes by combining three types of trees: cherry blossoms, maples, and pines.

At the top of each design scheme, visual representations of the three types of trees are shown: cherry blossoms with pink flowers, maples with orange leaves, and pines with green needles. Each design scheme features a distinct combination of these trees to create unique landscape effects.

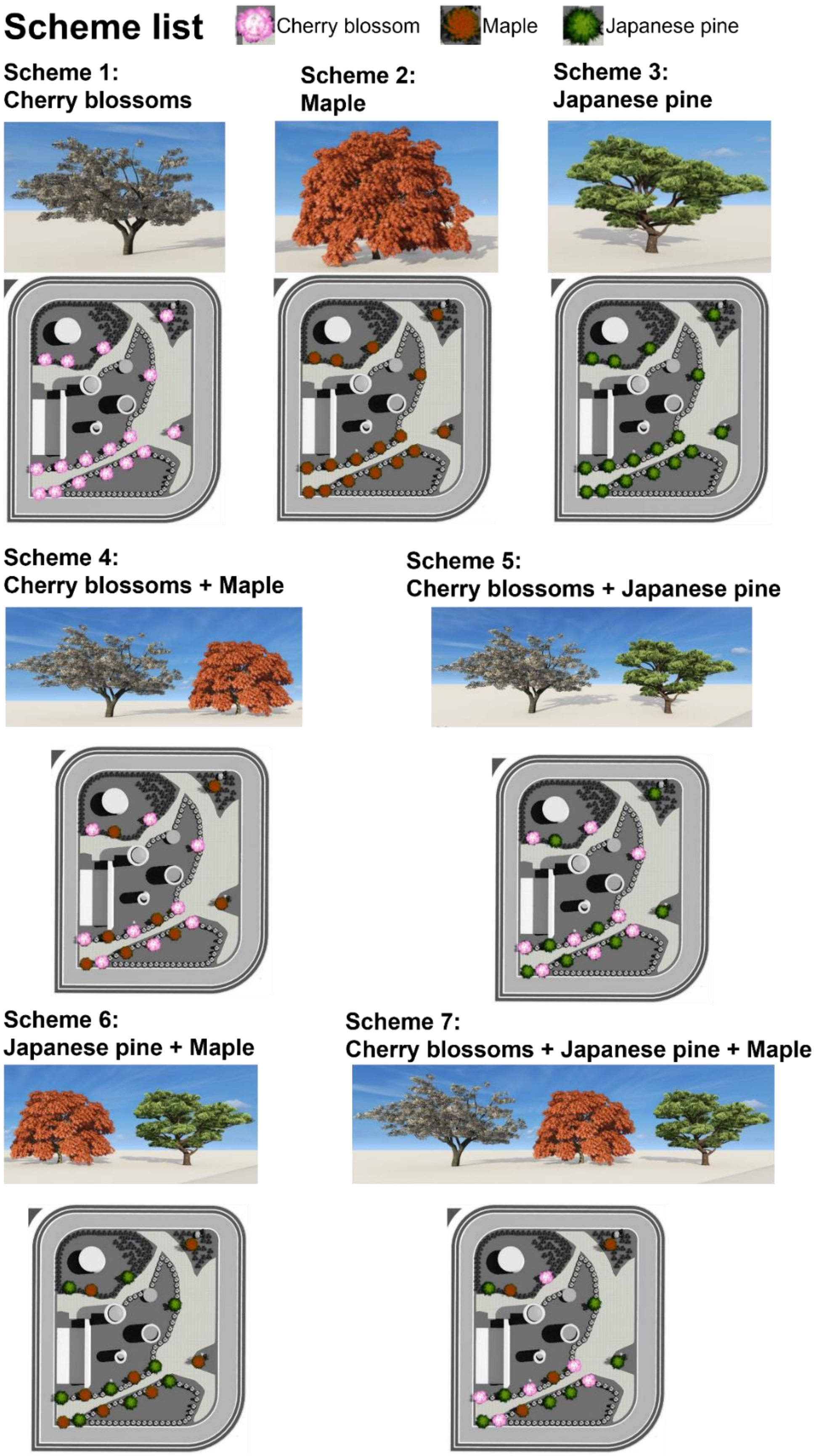


Figure 4.18 Seven planting design schemes for the virtual park

Each scheme is presented with a visual simulation of the trees in the upper half and a corresponding layout plan in the lower half. The schemes progress from single-species plantings to more complex combinations: the first uses cherry blossoms (pink markers), the second focuses on maples (orange markers), and the third highlights pines (green markers). The fourth scheme combines cherry blossoms and maples, demonstrating their visual interaction. The fifth mixes cherry blossoms with pines, while the sixth blends maples and pines. The seventh and final scheme incorporates all three tree types, offering the most diverse landscape composition. These vegetation configurations were implemented in a virtual park model, with fourteen key nodes selected for virtual camera placement to simulate tourist viewpoints. Using 3D rendering technology, high-quality scene images were generated for each scenario across all four seasons, enabling a comprehensive visual comparison of landscape effects across different configurations.

Figure 4.19 provides a detailed demonstration of Scheme 1, showcasing the seasonal changes observed from 14 camera locations in the virtual park. This figure captures the transformation of cherry blossoms through spring, summer, autumn, and winter, illustrating their distinct seasonal characteristics. Figure 4.20 demonstrates the seasonal changes observed at the same 14 camera positions when the trees are replaced with maples in Scheme 2. It is evident that the viewing seasons and visual effects of maple trees differ significantly from those of cherry trees.

Scheme 3, presented in Figure 4.21, exhibits pine trees, which show minimal changes across the four seasons. This consistency is observed at all 14 camera positions, highlighting the evergreen nature of pines. In Scheme 4, cherry trees and maple trees are combined. The arrangement is depicted in the plan view on the left side of Figure 4.22. More varied landscape changes are observed at the camera positions throughout the year, but it's important to note that both species will shed their leaves and become dormant in winter.

Figure 4.23 illustrates Scheme 5, which combines cherry trees and pine trees. Compared to Scheme 4, this combination results in less dramatic color changes in autumn. However, the presence of evergreen pine trees provides continuous landscape interest in winter, avoiding the bare appearance of deciduous trees.

Scheme 6, shown in Figure 4.24, combines pine trees and maple trees. While the colors may be relatively uniform in spring and summer, this combination produces a particularly striking landscape effect in autumn due to the maple's vibrant fall foliage contrasting with the evergreen pines. In Scheme 7, cherry trees, pine trees, and maple trees are integrated, as depicted in Figure 4.25. This comprehensive scheme ensures visual interest across all seasons: flowering trees in spring, lush foliage in summer, vibrant colors in autumn, and the enduring presence of non-deciduous pine trees in winter.

Scheme 1: Cherry blossoms

Figure 4.19 Seasonal transformation of vegetation in Scheme 1.

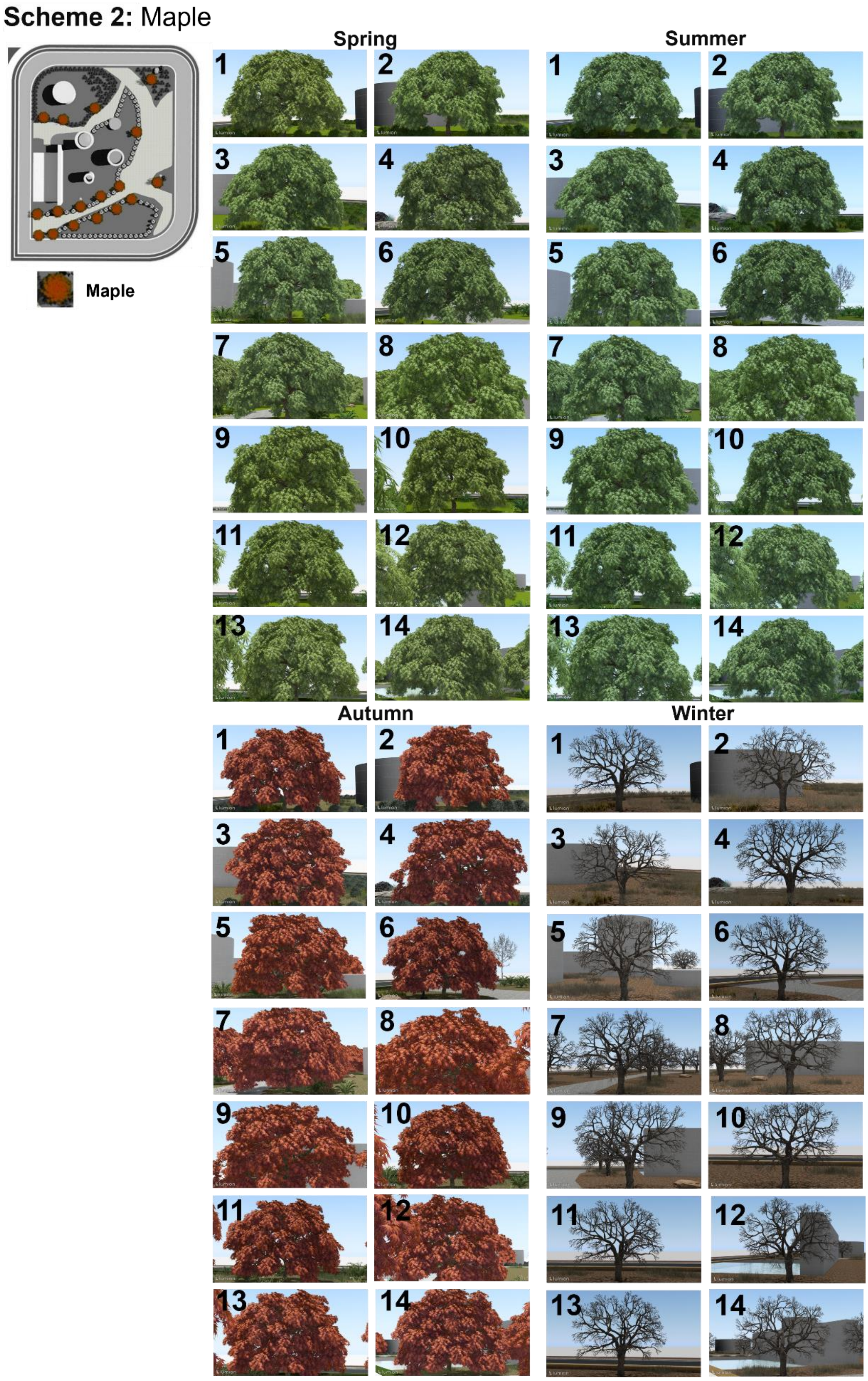


Figure 4.20 Seasonal transformation of vegetation in Scheme 2.

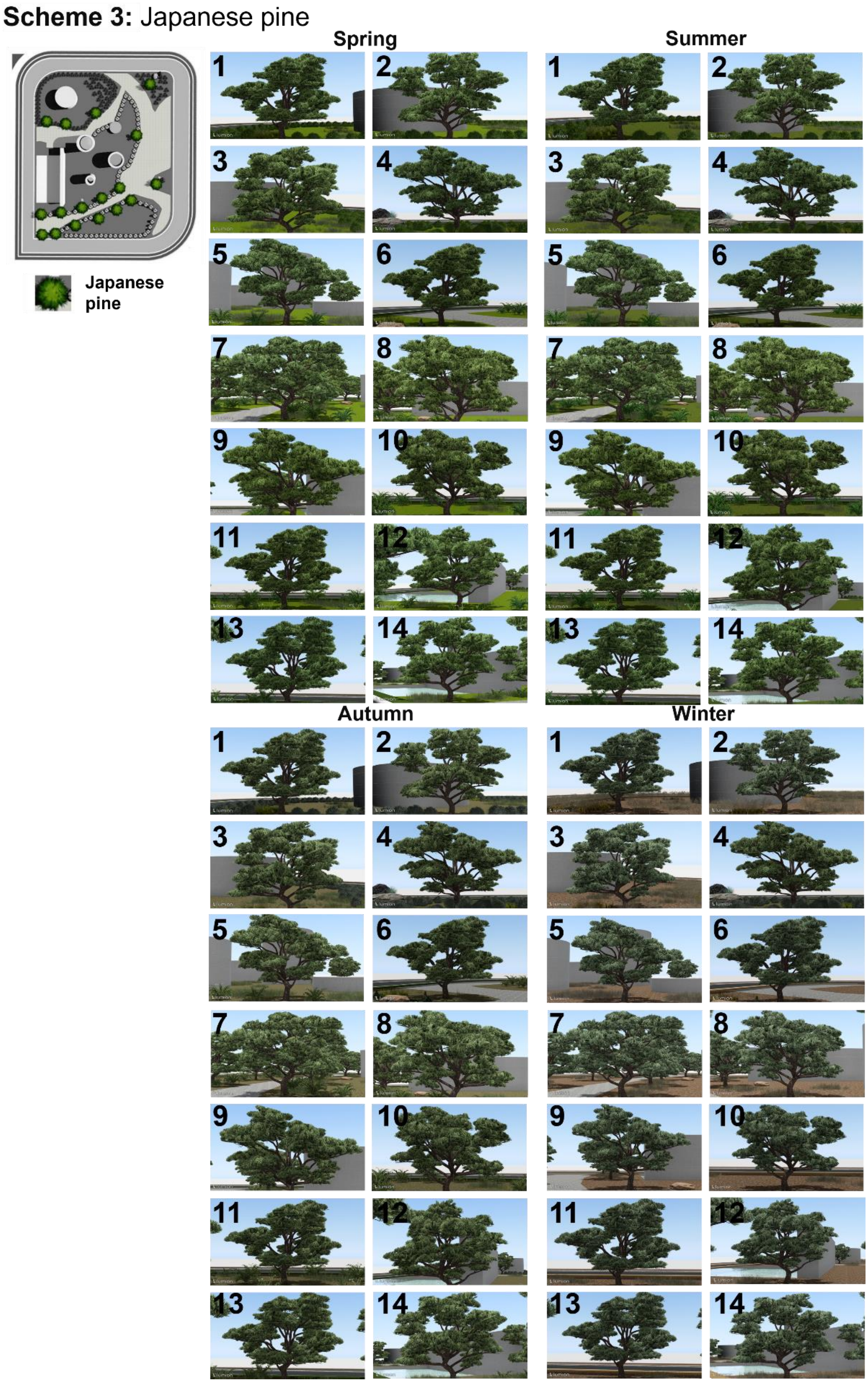


Figure 4.21 Seasonal transformation of vegetation in Scheme 3.

Figure 4.22 Seasonal transformation of vegetation in Scheme 4.

Figure 4.23 Seasonal transformation of vegetation in Scheme 5.

Figure 4.24 Seasonal transformation of vegetation in Scheme 6.

Scheme 7: Cherry blossom + Japanese pine + Maple

Figure 4.25 Seasonal transformation of vegetation in Scheme 7.

Figure 4.26 illustrates the S3PVI values for seven different planting design schemes across the four seasons. The S3PVI analysis of the seven planting schemes reveals insights into seasonal landscape dynamics, offering information for landscape design decisions. Schemes 1 (cherry blossoms) and 2 (maple) demonstrate high S3PVI values in spring and autumn respectively, highlighting the potential for single-species plantings to create powerful visual impacts during specific seasons. This approach can be ideal in certain contexts, such as creating iconic cherry blossoms avenues or maple viewing areas that offer visitors an intense, focused seasonal experience.

Scheme 3 (Japanese pine) exhibits the most consistent S3PVI values year-round, ranging from about 40% to 50%, underscoring the value of evergreen plants in maintaining landscape appeal throughout the seasons. Schemes 4, 5, and 6, which combine two tree types, achieve a better balance of visual interest across the year. For instance, Scheme 5 (cherry blossoms and Japanese pine) provides both spring spectacle and year-round greenery. Scheme 7, incorporating all three tree types, appears to offer the most comprehensive year-round appeal, with peaks in spring and autumn and sustained interest in summer and winter.

Comparing these schemes emphasizes key factors affecting park landscape visual features: plant species selection, spatial layout, and color matching. Scheme 7 demonstrates how combining cherry blossoms, maple, and Japanese pine can create diverse seasonal landscapes, offering rich spatial layers and color variations throughout the year. This multi-temporal visualization framework provides landscape designers with a tool to quantify and visualize S3PVI values across seasons for different planting schemes. It can help identify schemes that offer balanced, year-round appeal or those that might benefit from adjustments to enhance their off-peak season performance.

It's crucial to recognize that S3PVI analysis aims to provide objective metrics to support design decisions rather than replace designers' creativity and professional judgment. Its primary value lies in supporting and enhancing the subjective design process, enabling designers to make more informed decisions based on data. For example, designers might choose to maintain highly concentrated single-species plantings in certain areas for dramatic visual effect, while adopting more diverse plant configurations in others to ensure year-round landscape appeal. In this way, S3PVI analysis serves as a supportive tool, assisting designers in creating dynamic, seasonally balanced urban green spaces while meeting specific design objectives.

Figure 4.26 Seasonal S3PVI values for seven planting design schemes in the virtual park

4.4 Discussion

The multi-temporal urban green space vegetation visualization analysis framework constructed in this study demonstrates its capability to assess current conditions through the analysis of the Sanshikisaido case study, particularly suitable for initial assessments of brownfield projects. Quantitative indicators such as the S3PVI index and visual contribution rate reveal the differential roles of various plant types in shaping the visual effect of street landscapes. Based on the street view photo data of Sanshikisaido, this study analyzed the temporal changes of vegetation visual features with monthly precision. The virtual park project further validates the framework's potential in planning and designing vegetation layouts on 3D ground surfaces, offering a new perspective for urban green space planning and design decision-making. By creating virtual environments, this study simulated different vegetation configuration schemes and quantitatively evaluated these schemes using the S3PVI index. This method allows planners and designers to visualize and compare seasonal changes and visual effects of different design schemes before actual implementation.

4.4.1 Advantages and limitations of the multi-temporal urban green space vegetation visualization analysis framework

This research framework provides insights into the changes of urban vegetation visual characteristics by introducing a temporal dimension. However, it should be noted that the current framework mainly assesses the visual features of plants in their ornamental state. While this approach can capture the seasonal highlights of vegetation landscapes, it may not fully reflect the ecological value and visual contribution of plants during their non-ornamental periods.

Compared to traditional assessment methods such as GCR and GVI, S3PVI focuses on the plant type scale and can distinguish the visual contributions of different plant species. This feature gives S3PVI a unique advantage in assessing the diversity and seasonal changes of plant landscapes.

However, the research framework also has some limitations. The current plant dataset is limited in scale, which may affect the accuracy of the model in identifying diverse vegetation types. Moreover, existing technology still faces challenges in capturing plant features during periods of non-significant visual characteristics, which may lead to biases in identifying certain plant species in some seasons. For example, at certain stages of the growth cycle, such as cherry trees outside their flowering period, may be difficult to accurately identify and classify. These limitations suggest the need for further expansion of the dataset and improvement of identification algorithms to enhance the framework's applicability under different seasonal and diverse vegetation conditions.

The application of 3D Gaussian splatting for data processing and optimization in this framework addresses a critical issue in urban vegetation analysis: the inconsistency of street view images. Traditional methods often rely directly on street view images for calculations like GVI, which can lead to significant errors due to variations in camera angles, distances, and distortions. This framework's approach of using 3D Gaussian splatting to reconstruct 3D scenes from street view images, followed by the generation of standardized views, represents an advancement in data preprocessing for urban green space analysis.

The process, which involves SfM for initial 3D point cloud reconstruction and 3D Gaussian splatting for creating a continuous volumetric density representation, effectively mitigates issues of inconsistent perspectives and distortions in original street view images. By generating standardized views that simulate a consistent pedestrian perspective (at 1.6 meters height and with optimized distances based on plant height), the framework ensures a more accurate and reliable basis for S3PVI calculations.

However, it is important to note that while this method significantly improves data consistency, it also introduces computational complexity. Future research could focus on optimizing the efficiency of this process for large-scale urban analyses and exploring how this standardized approach might be integrated with or compared to other methods of urban vegetation assessment.

4.4.2 The significance of the S3PVI for quantitative evaluation

In constructing the multi-temporal visualization analysis framework, this chapter created a dataset containing 51 common urban environmental plants, providing a crucial data foundation for the development and validation of the S3PVI index. The construction of this dataset involved street view image collection, expert labeling, and multiple rounds of validation, ensuring the quality and representativeness of the data. These 51 plants cover common street trees, park plants, and seasonal ornamental plants, providing rich training and testing samples for the S3PVI index.

This dataset contributes significantly to the development of the S3PVI index, enabling it to identify and quantify the visual contributions of different plant types across seasons. For instance, it allows the system to distinguish the unique visual effects of cherry blossoms in spring and maple trees in autumn, thus providing quantitative basis for seasonal planning of urban green spaces.

However, there is still room for improvement in terms of plant species richness and annotation accuracy. Future research could consider expanding the dataset to include more plant species, especially those common in different climatic regions. Meanwhile, improving annotation precision, particularly for plants that are difficult to identify during non-ornamental periods, is also a focus for future work.

4.4.3 Application of the multi-temporal analysis framework in empirical case studies

Moreover, the current focus of the multi-temporal visualization analysis framework is primarily on the visual aesthetic features of vegetation, with insufficient attention to other green space functions such as ecological functions and recreational activities. Future studies should aim to integrate the S3PVI index with other ecological and recreational assessment indicators (Fix et al., 2018; Maclean et al., 2021) to achieve a comprehensive assessment of the multifunctional attributes of urban green spaces. Additionally, variations in topography, climate, and culture across different regions may influence people's preferences for plant landscapes (Hoyle et al., 2017). As the samples in this study are mainly from Japanese cities, the generalizability of the assessment results needs to be validated across a broader range. Future research should enhance the diversity of vegetation types and urban environments considered and incorporate influences from more diverse cultural backgrounds to improve the robustness and applicability of the assessment framework.

4.4.4 Multi-temporal vegetation visualization analysis supporting 4D vegetation landscape modeling

Compared to traditional field surveys and manual interpretation methods, multi-temporal visualization analysis offers significant advantages in efficiency and cost-effectiveness for data acquisition and automated processing. However, the accuracy of image processing and semantic segmentation still needs further improvement. Future research can explore integrating street view images with other data sources, such as high-resolution satellite images and LiDAR point cloud data, to provide more comprehensive and multi-scale information on urban green space vegetation.

Future research could explore integrating generative technology systems to achieve more complex 4D seasonal change simulations, further enhancing decision support capabilities in planning and design processes. This integration could allow for more accurate predictions of how vegetation will change over time, taking into account factors such as growth rates, seasonal variations, and environmental influences. Such advancements would provide planners and designers with a more dynamic and realistic view of their proposed green spaces over extended periods.

Additionally, the development of interactive design tools could enable planners to directly design vegetation layouts on 3D ground surfaces, creating a more intuitive and interactive planning process. These tools could incorporate real-time feedback on factors such as visual impact, ecological value, and maintenance requirements, allowing for more informed decision-making. The use of virtual reality (VR) or mixed reality (MR) technologies could further enhance this process, allowing stakeholders to virtually walk through proposed designs and experience them from multiple perspectives.

By coupling S3PVI spatiotemporal change features with 4D modeling tools, a detailed depiction of the growth and development process of garden plants can be achieved. This could lead to the creation of an immersive garden landscape design platform, where planners, designers, managers, and even the public could collaboratively participate in the entire process of urban green space creation. Such a platform would not only improve the quality of designs but also increase public engagement and acceptance of urban green space projects.

4.4.5 *Improvement directions and application extensions*

Looking towards the future, the S3PVI analysis framework has the potential to extend beyond its current focus on 2D image analysis. By expanding into 3D space and multi-sensory dimensions, future iterations of the framework could explore the correlations between vegetation visual features and other elements of the urban landscape experience. This could include factors such as garden spatial structure, recreational activities, soundscape, and even olfactory aspects, providing a more holistic understanding of how vegetation contributes to high-quality urban environments.

The integration of these multi-dimensional analyses with the S3PVI framework could lead to a new paradigm in urban green space governance driven by digital twins. In this scenario, vegetation landscapes' assessment, simulation, perception, and feedback would form a dynamic closed loop. Real-world data would continuously inform and refine digital models, while these models would in turn guide real-world interventions and management strategies. This approach could revolutionize how urban green spaces are planned, designed, and managed, allowing for more responsive, adaptive, and effective strategies.

As these technologies and methodologies continue to evolve, it will be crucial to validate their effectiveness and practicality through real-world applications and case studies. Future research should focus on expanding plant datasets, improving identification algorithms, and exploring integration with other urban planning and management tools. By doing so, the value and reliability of this framework in practical applications can be enhanced, ultimately contributing to the creation of more sustainable, resilient, and enjoyable urban green spaces.

4.5 Summary of this chapter

Chapter 4 presents a multi-temporal urban green space vegetation visualization analysis framework, applied to the Sanshikisaido case study and a virtual park project. The framework introduces the S3PVI index for quantifying the visual contribution of different plant types in urban landscapes across seasons.

The study employs 3D reconstruction from street view images, deep learning-based image classification, and semantic segmentation to analyze vegetation characteristics. A dataset of 51 common urban plants supports the development and validation of the S3PVI index.

The Sanshikisaido case study demonstrates the framework's ability to capture seasonal changes in vegetation visibility, providing insights into the visual contributions of various plant species throughout the year. The virtual park project illustrates the potential of this approach in supporting decision-making for urban green space planning and design.

The chapter discusses the framework's capabilities compared to traditional assessment methods, noting its ability to differentiate between plant types and capture temporal variations. It also acknowledges limitations, such as the focus on plants in their ornamental states and challenges in identifying plants during non-distinctive visual periods.

Future research directions are outlined, including the expansion of the plant dataset, improvement of identification algorithms, and integration with other urban planning tools. The potential for extending the framework into 3D space and multi-sensory dimensions is considered, along with the possibility of developing more interactive design tools.

The chapter sets the stage for Chapter 5, which will explore 4D vegetation landscape modeling, integrating technologies such as NeRF and Stable Diffusion to create time-evolving models of urban vegetation.

The conclusion considers the implications of this research for urban green space governance, suggesting that multi-dimensional analyses could contribute to more responsive and effective strategies for creating sustainable urban environments. The framework presented in this chapter aims to advance urban vegetation assessment methods and contribute to the evolving field of urban green space management.

Chapter 5
Plant landscape modeling: integrating dynamics and techniques

5.1 Overview

Chapter 4 introduced a multi-temporal urban green space vegetation visualization framework and the S3PVI to characterize the visual importance of urban green spaces at the species level. The S3PVI quantifies the visual characteristics of different plant species across spatial and temporal dimensions, providing a foundation for understanding the dynamic features of urban vegetation. By focusing on the plant scale and capturing temporal changes in visual importance, S3PVI facilitates 4D time-varying landscape modeling.

Inspired by the findings in Chapter 4, this chapter explores the integration of temporal dynamics with advanced imaging technologies to develop a 4D modeling approach for plant landscapes. The aim is to create immersive dynamic representations of urban green spaces, supporting intuitive and flexible planning and design. The multi-temporal vegetation segmentation images and the spatiotemporal variation characteristics of S3PVI obtained in Chapter 4 guide the generation of realistic seasonal changes in the 4D modeling process. In addition to using this data to guide image generation, S3PVI values are incorporated into the Stable Diffusion model as inputs to describe the visualization degree of visible plants in the scene. This integration further enhances the accuracy and realism of the generated plant landscapes by providing quantitative information on the visual prominence of different plant species.

This modeling approach employs NeRF and Stable Diffusion techniques to visualize plant growth and seasonal changes. The integration of low-rank adaptation (LoRA) (E. J. Hu et al., 2021) and ControlNet (L. Zhang et al., 2023) enhances computational efficiency and ensures spatial consistency, enabling the simulation of plant development over time. The prior information provided by S3PVI, including multi-temporal vegetation segmentation images, spatiotemporal variation characteristics, and quantitative descriptions of plant visibility, contributes to improving the accuracy and diversity of the generated images, facilitating the integration between multi-temporal analysis and 4D modeling.

These tools have the potential to provide urban planners and environmental scientists with additional means to observe, predict, and plan urban vegetation growth patterns. If proven effective through further research and validation, they could contribute to supporting sustainable and responsive urban green space management, as suggested by Nitoslawski et al., (2019). The combination of multi-temporal visualization and 4D modeling broadens the scope of urban green space research, providing a comprehensive approach to understanding and managing urban vegetation dynamics.

This chapter presents the technical framework for constructing 4D plant models, discusses innovative methods for achieving these dynamic simulations, and showcases a practical application case study, an actual park streetscape. Extending the static models discussed in previous chapters to include time-based dynamics, it provides new perspectives on the interaction between urban environments and their green spaces. The integration of plant lifecycle data into urban ecological planning and design could potentially play a significant role in improving the effectiveness of these processes.

The combination of multi-temporal visualization and 4D modeling contributes to advancing urban green space research, supplementing traditional methods. By leveraging advanced technologies and building on the foundation laid in Chapter 4, this framework enhances the understanding of urban green spaces and provides tools for sustainable urban development and ecological planning. Incorporating temporal dynamics into urban vegetation visualization offers a potentially comprehensive approach to visualizing and managing urban vegetation holistically and interactively. This may help bridge the gap between static and dynamic representations of urban green spaces.

5.2 Proposed framework

5.2.1 Technical framework construction

The backbone of the proposed 4D plant modeling framework consists of two key components: NeRF with multiresolution hash encoding for learning 3D scene representations and stable diffusion for synthesizing time-varying appearances. Choosing NeRF with multiresolution hash encoding was driven by its capability for faster training and rendering of 3D scenes. Its real-time performance makes it highly suitable for interactive applications like

VR, enabling seamless exploration and manipulation of 3D models. In contrast, the 3D Gaussian splatting method, which performed well in Chapter 4, lags behind NeRF in these aspects.

Table 5.1 compares the performance of NeRF (multiresolution hash encoding) and 3D Gaussian splatting across different metrics, including the SSIM, PSNR, Frames Per Second (FPS), and training time. While 3D Gaussian splatting achieves higher SSIM, PSNR and FPS scores, indicating better image quality, NeRF (multiresolution hash encoding) excels in terms of training time, making it more suitable for real-time applications and efficient scene processing.

Table 5.1 Performance comparison between NeRF and 3D Gaussian splatting

Metric	NeRF **(multiresolution hash encoding)**	**3D Gaussian splatting**
SSIM	0.72	0.83
PSNR	24.92	26.91
FPS	9.00	137.30
Training Time	7.5 min	38.3 min

To complement the strengths of NeRF with multiresolution hash encoding in capturing 3D scene representations, Stable Diffusion was selected for synthesizing time-varying appearances in the proposed 4D plant modeling framework. Stable Diffusion, a state-of-the-art text-to-image generation model, leverages the diffusion model paradigm to convert random noise into target images through iterative forward and reverse diffusion processes. By incorporating text embeddings as additional inputs, Stable Diffusion enables a conditional control mechanism that guides the image generation process towards the desired seasonal characteristics specified by the text prompts.

Moreover, Stable Diffusion's latent space representation and attention mechanism allow it to establish semantic associations between textual descriptions and image content, ensuring the generated vegetation images align with the given seasonal features. The model's rich training data and strong generative capabilities enable it to accommodate the representation requirements of various vegetation types and seasonal variations. The open-source nature and active community support of Stable Diffusion provide flexibility and extensive reference resources for its application in this research. By integrating Stable Diffusion with the 3D scene representations learned by NeRF, the proposed framework can introduce realistic seasonal changes to vegetation while maintaining spatial consistency, ultimately achieving dynamic and photorealistic 4D plant model construction.

5.2.2 Model generation pipeline

The proposed method adopts a cascaded approach to generate 4D plant models in two stages. In the first stage, NeRF is trained for each input scene to learn a compact 3D representation of its geometry and appearance. The training process involves sampling 5D coordinates from different camera poses, which include three-dimensional spatial coordinates (x, y, z) and two-dimensional viewing direction information. These coordinates are then encoded into high-dimensional feature representations using specialized encoding techniques, such as hash encoding or quadtree encoding. A multi-layer perceptron (MLP) network processes these high-dimensional features to predict colors and densities at each point. The results are rendered using volume rendering techniques, which combine the predicted colors and densities along the light paths to generate the final images. The MLP network is optimized using a photometric reconstruction loss, which minimizes the difference between the rendered images and the ground-truth images to ensure high fidelity.

In the second stage, the scene photos, along with text prompts specifying the desired seasonal variations, are input into the Stable Diffusion pipeline. It is worth noting that this stage makes full use of the multi-temporal vegetation segmentation images and S3PVI values obtained in Chapter 4 to guide the synthesis of realistic vegetation images. Stable Diffusion utilizes these inputs to manipulate the latent representation of the images. The process begins with the scene photos, and by gradually denoising and decoding the refined latent code, Stable Diffusion synthesizes plant images that exhibit the targeted seasonal characteristics while preserving the original scene content. The integration of text prompts helps guide the diffusion model to introduce the specified seasonal changes. To ensure spatial consistency, the initial renderings from NeRF are used as a baseline, upon which Stable Diffusion builds to introduce appearance changes. This approach leverages the strengths of both NeRF and Stable Diffusion to generate realistic and seasonally varied 4D plant models.

To further improve the visual quality and geometric consistency of the generated images, the proposed method integrates LoRA and ControlNet into the pipeline.

LoRA is a parameter-efficient fine-tuning technique that adapts pre-trained weights through low-rank decomposition, enabling fast and effective fine-tuning for downstream tasks. In this study, LoRA is applied to pre-process the input images, emphasizing key structures such as trunks and branches.

The LoRA training process is crucial, enabling Stable Diffusion to comprehend the meaning of S3PVI values and generate images with corresponding vegetation visibility. The steps include: (1) collecting over 50 medium-resolution images of each plant species across different seasons; (2) standardizing all images to 512x512 pixel resolution to ensure consistent processing; (3) performing automatic image annotation using the BLIP neural network, carefully labeling each image according to its corresponding season; (4) employing the Kohya_ss framework (Bmaltais, n.d.) for LoRA training, which has been verified for compatibility and performance with Stable Diffusion. Finally, the result graphs of each LoRA model are analyzed, focusing on the model weight effectiveness and performance indicators, rather than relying solely on reducing the loss. Through this process, the LoRA models can effectively learn the seasonal characteristics of plants, preparing them for the subsequent image synthesis stage.

During the training data preparation phase, each vegetation image is combined with its corresponding S3PVI value and image description text. This allows the LoRA model to establish a mapping between image visual features, S3PVI quantitative indicators, and textual semantic descriptions. Once trained, the Stable Diffusion model, equipped with the LoRA plugin, can understand the vegetation visibility level represented by a given S3PVI value and reflect it in the image generation process, ensuring that the generated vegetation images have visibility rates matching the input S3PVI values.

Furthermore, the multi-temporal vegetation segmentation images obtained in Chapter 4 are applied to ControlNet's semantic segmentation control. By utilizing pre-generated segmentation masks, ControlNet can explicitly define the spatial scope of Stable Diffusion's image modifications, focusing on the vegetation regions and avoiding interference with the background. The segmentation masks also provide Stable Diffusion with a reference for the spatial layout of the vegetation, ensuring that the generated images maintain structural consistency with the original scene.

By establishing the association between S3PVI and vegetation visibility levels through LoRA training and integrating multi-temporal vegetation segmentation information into ControlNet, Stable Diffusion can more accurately and elaborately represent the visual features of vegetation across different seasons. The S3PVI prior knowledge learned by LoRA enables the generated images to quantitatively reflect the visibility of vegetation, while the segmentation-guided ControlNet further optimizes the spatial structure expression of the vegetation. The synergy between the two enhances the 4D modeling method's ability to capture the temporal and spatial variation patterns of vegetation simultaneously, reconstructing the dynamic process of plant landscapes with higher realism and detail richness.

This modeling approach, which combines S3PVI and semantic segmentation with LoRA and ControlNet, fully utilizes the spatiotemporal feature information of vegetation revealed by the multi-temporal visualization analysis in Chapter 4, transforming it into prior knowledge that Stable Diffusion can understand and leverage. This greatly improves the accuracy and expressiveness of 4D scene construction. It represents a new perspective, achieving full-time, high-fidelity digitalization and simulation of urban green spaces through the integration of multi-dimensional data and methods.

5.2.3 The main process of the system

The process involves several key steps, as shown in Figure 5.1:

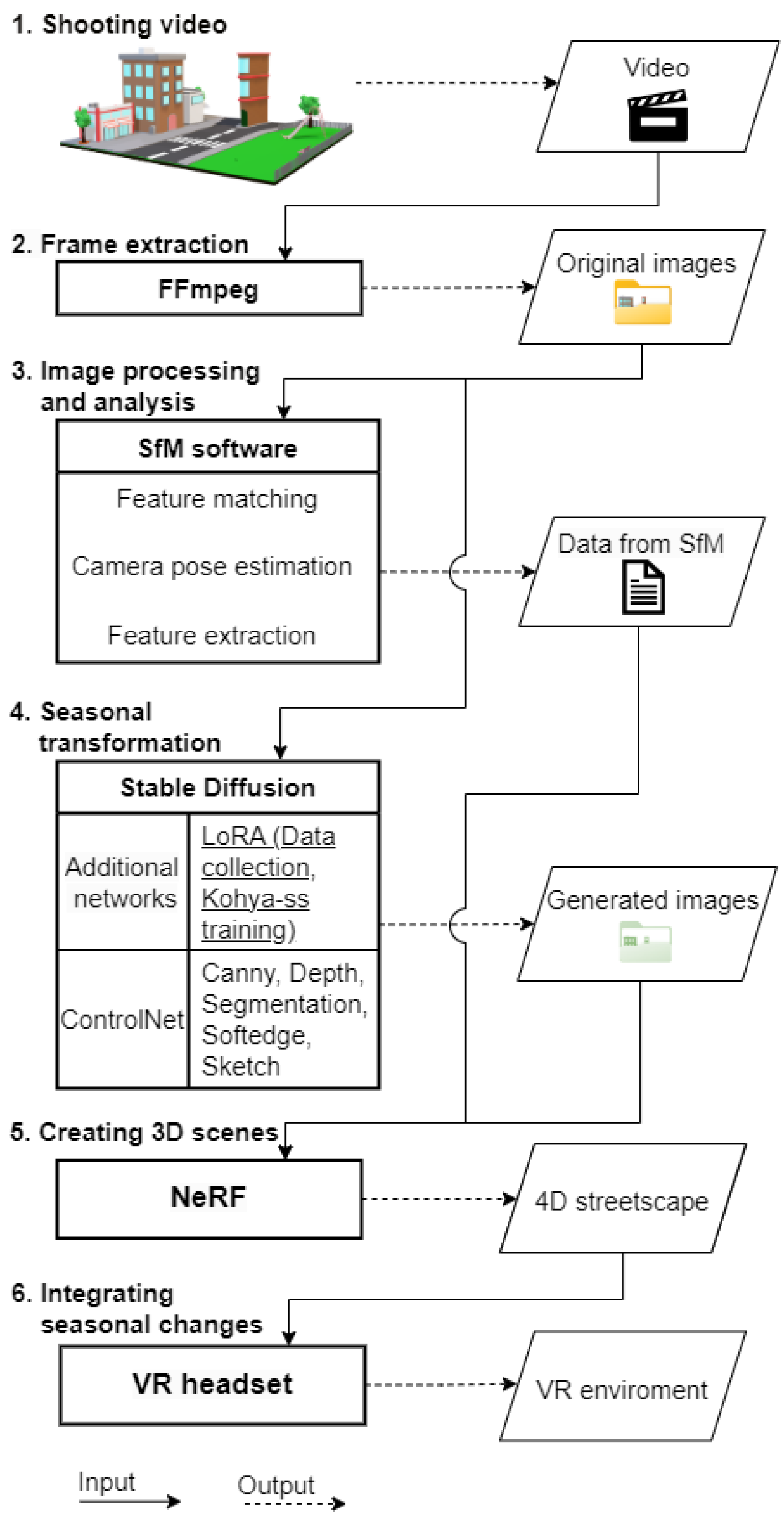


Figure 5.1 4D generation system for plant landscape modeling.

1) Video capture: A comprehensive video of the intended plant is recorded to ensure full coverage from every angle. Plants, within the scope of video production, can be categorized as either real or virtual. This is achieved by filming over a full 360-degree rotation at a consistent speed.

2) Frame extraction: The video is fragmented into frames using a multimedia processing tool (*FFmpeg Developers,* 2016). The count of frames obtained is reliant on the length of the video, intending to secure over 200 frames to facilitate comprehensive scrutiny.

3) Image processing and analysis: Preliminary analysis for 3D reconstruction involves crucial steps such as feature extraction, feature matching, and camera pose estimation. Tools such as COLMAP can be used for this purpose (Schonberger & Frahm, 2016).

4) Seasonal transformation via Stable Diffusion: Frames are input into Stable Diffusion to generate images of the plant in different seasonal states. Techniques such as LoRA are utilized for accurate seasonal depiction, while ControlNet maintains the plant's structural integrity.

5) Creating 3D scenes with Instant-ngp, the integrated framework for NeRF implementation with multiresolution hash encoding: These frames, along with the data derived from COLMAP, are input into Instant-ngp. This step generates a realistic 3D rendition of the plant, which can be explored in VR, offering real-time interaction and the ability to modify various aspects such as lighting and viewpoint.

6) Integrating seasonal changes in 3D Scenes: The seasonally transformed images, combined with the data from the COLMAP analysis, are then processed through Instant-ngp. This step creates a detailed 3D scene reflecting the plant in various seasons.

5.2.4 Model evaluation metrics

To evaluate the quality and realism of the generated 4D plant models, several metrics are employed:

1) PSNR measures the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Higher PSNR values indicate better image quality and less distortion compared to the ground truth.

2) SSIM assesses the perceived quality of digital images and videos by quantifying the similarity between two images based on luminance, contrast, and structure. SSIM values range from -1 to 1, with higher values indicating greater structural similarity to the reference image (Hore & Ziou, 2010).

3) Learned perceptual image patch similarity (LPIPS) is a learned metric that measures the perceptual similarity between two images using deep features extracted from a pre-trained CNN. Lower LPIPS scores suggest higher perceptual similarity to the ground truth (R. Zhang et al., 2018).

These metrics provide a comprehensive assessment of the generated 4D plant models, considering both low-level image quality and high-level perceptual similarity. By evaluating the models using these metrics, the effectiveness of the proposed framework can be quantitatively validated.

5.3 Experiments and results

5.3.1 Experimental setup

To demonstrate the effectiveness of the proposed 4D plant modeling framework, experiments were conducted using a dataset of virtual plants. The use of virtual plants allows for controlled testing and refinement of the approach, as obtaining comprehensive data from real plants across all seasons can be challenging. Table 5.2 provides a detailed summary of the equipment, software, and model versions used in this study, along with the specific parameters employed for training the LoRA model.

Table 5.2 Equipment, software, and parameters used.

Category	Equipment /Software	Specifications/Version	Parameters for LoRA training	Value
PC specifications	CPU	Intel Core i5-11400	LoRA type	Standard
	GPU	NVIDIA GeForce RTX 3060 Ti	LR scheduler	Cosine_ with_ restarts
	RAM	DIMM 16 GB	LR Warmup (% of steps)	10
Analytical model & Software	FFmpeg	Version: 5.1.2	Optimizer	AdamW8 bit
	COLMAP	Version: 3.7-windows-no-cuda	Max resolution	512x512
	Instant-ngp	RTX-3000-and-4000	Network Rank and Alpha	128, 128
	Stable Diffusion	Version ID: baf6946	Total steps	2000
	Equipment /Software	**Specifications/Version**	**Parameters for LoRA training**	**Value**
	Additional networks	Version ID: e9f3d62	Train Batch Size	1
	ControlNet	Version ID: 3011ff6	Epoch	1
	Kohya_ss	Version 2.0	Regularization factor	1
			Mixed precision	fp16

5.3.2 Data and target plant selection

For this study, a summer maple tree was selected as the target plant for photographic data collection to train the LoRA algorithm. The distinctive leaves and branching patterns of the summer maple provide clear and consistent data points that are ideal for algorithmic analysis. The lush summer foliage ensures seasonal consistency, providing a rich dataset. The tree's common presence and photogenic nature facilitate accessibility and high-quality image capture, which are critical for effective algorithm training. These factors contribute to a robust and diverse dataset, improving the accuracy and efficiency of the LoRA algorithm, which is essential for generating precise and varied simulated models.

Using the LoRA training method outlined in Section 5.2.2, a dataset of 50 photographs featuring maple plant models in the target season was compiled. These images were then processed using the multi-temporal urban green space vegetation visualization framework to generate segmented images and calculate their corresponding S3PVI values. The S3PVI values for each image were subsequently recorded in dedicated text files. To train the LoRA model, the Kohya_ss framework was employed, utilizing the segmented images as input data. Additionally, these segmented images were fed into Controlnet to serve as a reference for refining certain constraints within the model. Finally, this information will be used in the Stable diffusion.

Figure 5.2 illustrates a comprehensive process of data collection, image processing, model training, and image generation, using a summer maple tree as an example. The workflow is divided into four main steps, each contributing to the overall goal of analyzing and recreating urban green spaces:

1) Collecting images: This initial step involves capturing target seasonal images using cameras. The focus is on obtaining high-quality, representative maple trees images.

2) Processing images: The collected images are then fed into a "Multi-temporal urban green space vegetation visualization framework". This sophisticated system processes the raw images to extract valuable data. It generates S3PVI values, which quantify vegetation. Additionally, it produces segmented images that highlight specific areas of interest. These segmented images are crucial as they focus the algorithm's attention on key visual elements that contribute to the tree's unique appearance.

3) Training the model: The processed data is used to train a model named "Kohya_ss". This step likely involves machine learning techniques to teach the model to recognize and understand the characteristics of maple trees. The training process incorporates both the S3PVI values and the segmented images, allowing the model to learn from both quantitative data and visual features.

4) Generating images: The final step utilizes the trained LoRA model in conjunction with original images. Through Stable Diffusion technology, new images are generated. This

process aims to create realistic, detailed images of urban green spaces that reflect the learned characteristics of the maple trees and other vegetation.

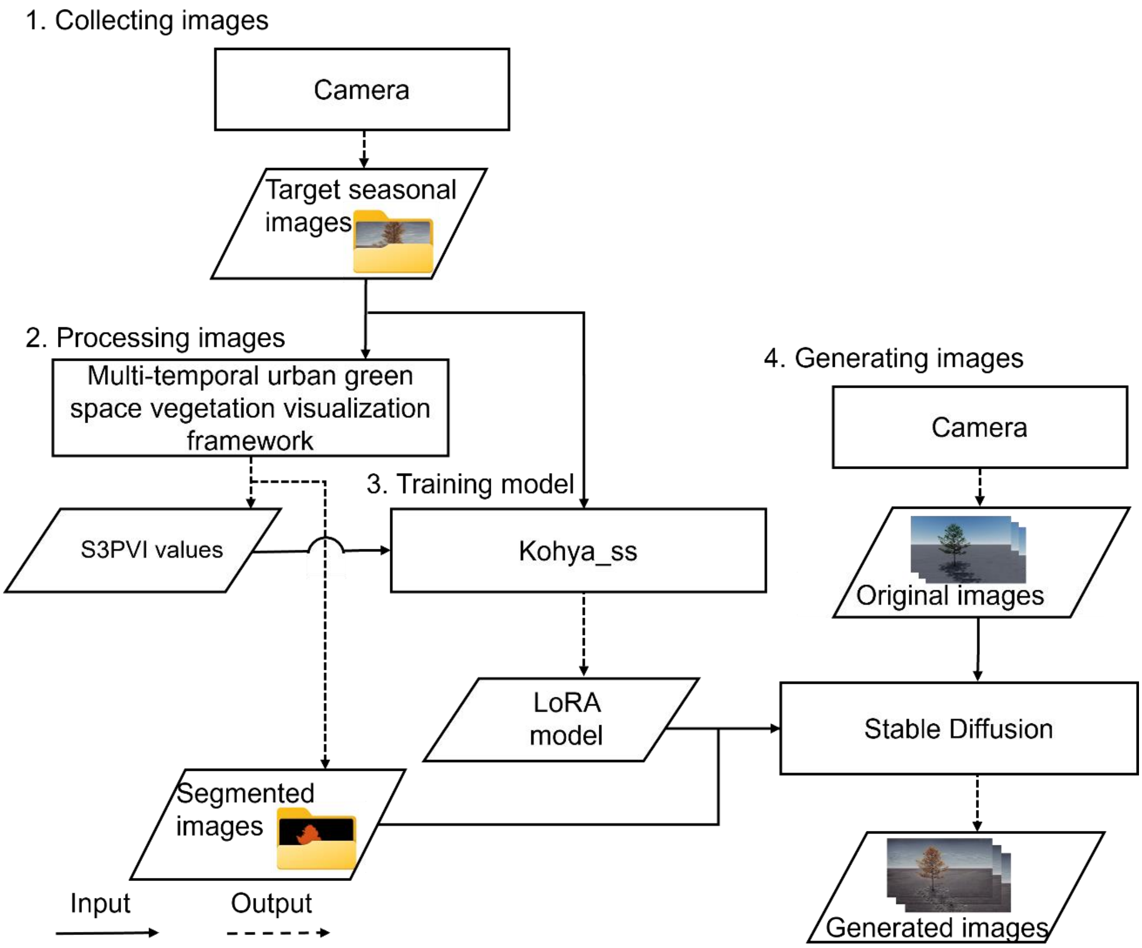


Figure 5.2 Workflow for training LoRA model using multi-temporal vegetation visualization.

5.3.3 Seasonal variation results

Figure 5.3 demonstrates the effectiveness of the LoRA model within the Stable Diffusion framework, showcasing a tree in different seasons and angles. The image is divided into four rows, representing spring, autumn, late autumn, and winter.

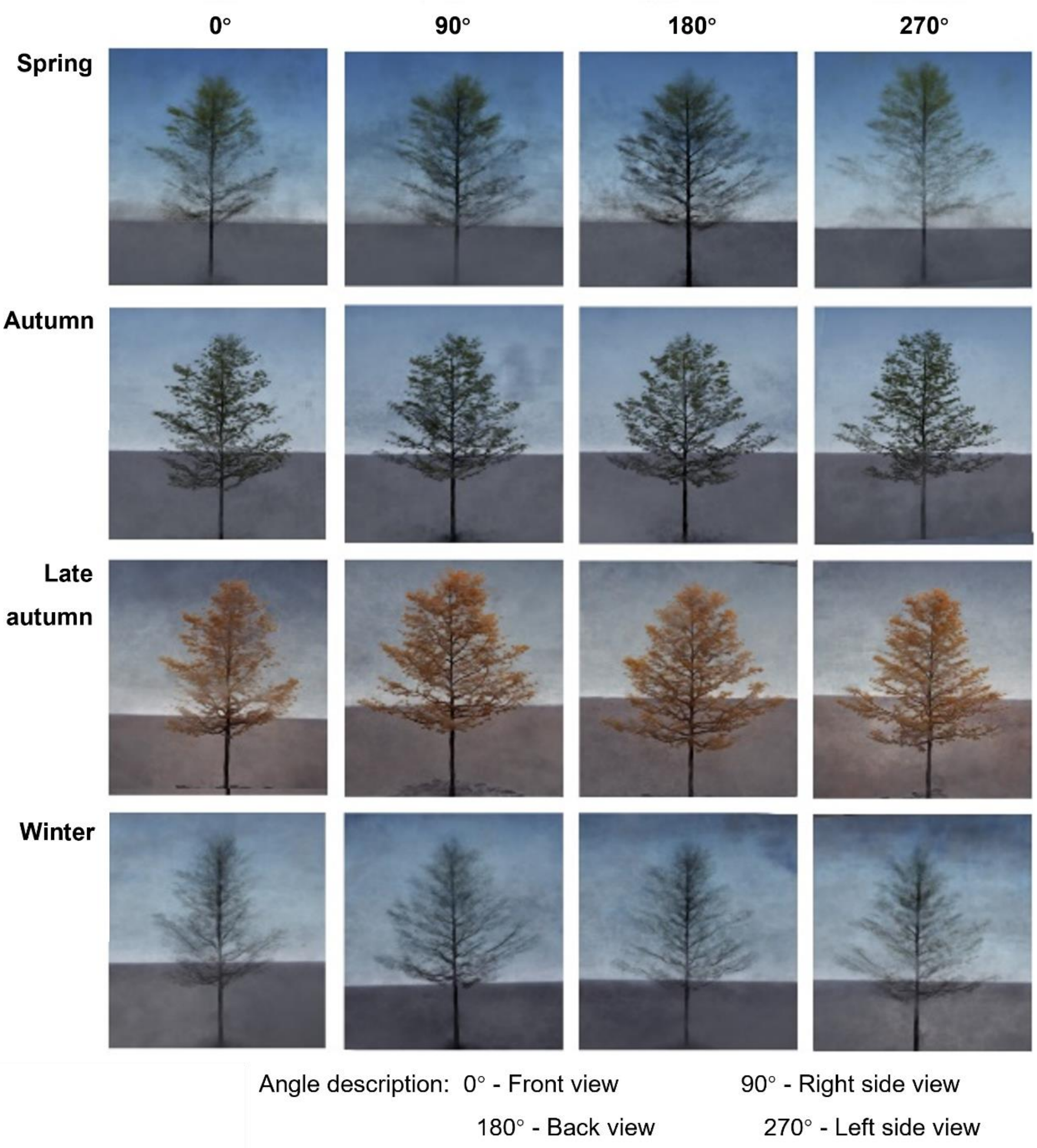


Figure 5.3 The state of the target plant in the VR environment in all seasons.

Each row contains four images labeled 0°, 90°, 180°, and 270°, displaying different perspectives as the tree rotates 360 degrees. The 0° view (leftmost) shows the front view, 90° the right side view, 180° the back view, and 270° (rightmost) the left side view. The integration of LoRA and ControlNet ensures that the generated images maintain spatial consistency and adhere to the original plant's edge structure while introducing realistic seasonal variations. These changes are reflected in the color and density of the foliage: sparse leaves in spring, full foliage in autumn, orange-brown leaves in late autumn, and bare branches in winter. By presenting different angles and seasons, the image comprehensively illustrates the tree's growth cycle changes, highlighting the model's ability to generate seasonal variations while maintaining spatial structure. Table 5.3 lists all the parameters applied, further illustrating the seasonal variations of the plant imagery.

Table 5.3 Parameters used in Stable Diffusion for seasonal variations.

Parameter	Value
Steps	26
Denoising strength	0.33
CFG scale	10
Seed	3373133097
Spring prompt	green leaf, bare tree, LoRA:winter:0.2
Autumn prompt	yellow leaf, LoRA:maple_autu:0.6
Late autumn prompt	maple tree, LoRA:maple_autu:0.6
Winter prompt	bare tree, LoRA:winter:0.6

Negative prompt: ground, background, trunks, branches, tree roots, small trees, people, other trees, classifier for paintings, etc, background, building, trunks, branches, small trees, grass, mountains, lake, sloping land, kkw-Autumn

ControlNet 0 Module: canny, Model: None, Weight: 1.2, Resize Mode: Resize and Fill, Low Vram: False, Processor Res: 512, Threshold A: 100, Threshold B: 200, Guidance Start: 0, Guidance End: 1, Pixel Perfect: False, Control Mode: ControlNet is more important

ControlNet 1 Module: depth_midas, Model: None, Weight: 1, Resize Mode: Crop and Resize, Low Vram: False, Processor Res: 512, Guidance Start: 0, Guidance End: 1, Pixel Perfect: False, Control Mode: Balanced

ControlNet 2 Module: seg_ofade20k, Model: None, Weight: 1.25, Resize Mode: Crop and Resize, Low Vram: False, Processor Res: 512, Guidance Start: 0, Guidance End: 1, Pixel Perfect: False, Control Mode: ControlNet is more important

5.3.4 ControlNet model evaluation

Figure 5.4 presents the results obtained from the three ControlNet models employed in this study. The left image demonstrates the Canny model's ability to capture the edges of the tree by converting the plant images into line drawings while maintaining compositional consistency. The center image shows the depth estimation model's accurate reproduction of the three-dimensional structure of the tree by extracting depth maps and reconstructing the spatial layout. The right image illustrates the semantic segmentation model's capability to segment the image into tree and background pixels.

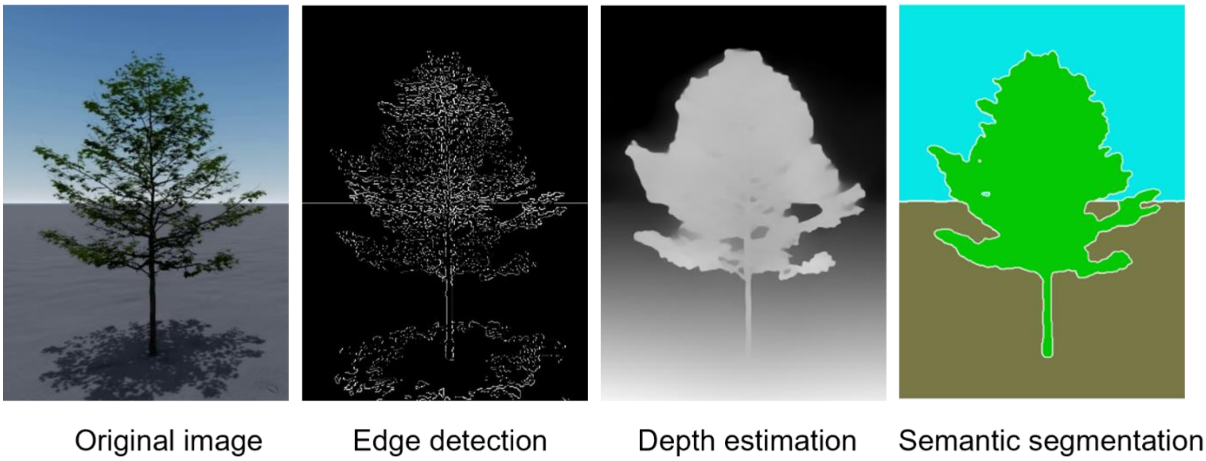


Figure 5.4 Comparative results of ControlNet models for edge detection, depth estimation, and semantic segmentation.

5.3.5 Image quality evaluation

To assess the quality of the generated plant scenes, multiple evaluation metrics were used, including PSNR, SSIM, and LPIPS. These metrics provide a comprehensive evaluation of the perceptual errors, visual changes, and perceptual similarity between the generated images and the reference frames. The evaluation process begins with Instant-ngp converting the 2D images into a 3D scene, focusing on light and color details. The reference frames serve as a benchmark for comparison, and Instant-ngp aligns the generated frames with these reference frames. PSNR measures the peak errors, with higher values indicating better quality. SSIM assesses the visual similarity, with values closer to 1 representing higher similarity. LPIPS, which utilizes deep learning, prefers lower scores for a closer resemblance to the original images. Random frame selection ensures an unbiased evaluation across different scenes.

Table 5.4 presents the comparative analysis of the image quality metrics for the generated plant scenes across different seasons. The high PSNR and SSIM scores, along with the low LPIPS scores, demonstrate the high quality of the images produced by Instant-ngp, showcasing its ability to accurately depict seasonal changes in the VR environment.

Table 5.4 Comparative analysis of image quality metrics.

	Spring	Autumn	Late autumn	Winter
PSNR	31.512	39.468	39.455	29.366
SSIM	0.952	0.965	0.848	0.733
LPIPS	0.457	0.398	0.432	0.589

5.3.6 Seasonal variation results for streetscapes

Figure 5.5 presents a comparative analysis of the generated summer streetscape park and the original winter streetscape, aiming to evaluate the effectiveness of the proposed system in simulating seasonal changes. The figure comprises three sub-images: (a) displaying the original winter image, (b) showing the system-generated summer image, and (c) presenting a real summer scene.

Through comparative analysis, several key differences can be observed. Trees in the real summer image appear taller, with more abundant leaves and richer colors, highlighting areas for improvement in the system's simulation of seasonal vegetation changes. Simultaneously, the flower beds in the real scene show less dense ground vegetation, reflecting the complex influence of various factors on plant growth in real environments, subtle environmental variables that have not been fully accounted for by the current generative model. Furthermore, the real summer photograph demonstrates a brighter and sunnier environment, underscoring the limitation of this technology, which primarily focuses on vegetation changes without adjusting overall lighting conditions.

Despite these differences, the generated summer scene overall demonstrates a satisfactory seasonal transition effect. These observations provide valuable insights for future research, indicating potential directions for further enhancing the realism and accuracy of seasonal transformation models.

These results demonstrate the potential of the proposed framework in generating 4D plant models that aim to capture both the spatial structure and temporal dynamics of plants. The integration of NeRF, Stable Diffusion, LoRA, and ControlNet enables the creation of plant visualizations across different seasons, offering a new approach to urban landscape modeling and analysis. While there is room for improvement in certain details, such as more accurate simulation of tree growth and environmental lighting changes, the overall approach demonstrates some effectiveness in seasonal transition. This research provides insights and possible directions for further enhancing the realism and accuracy of seasonal transformation models in the future.

(a) Original winter scene (b) Generated summer scene (c) Actual summer scene

Figure 5.5 Comparative analysis of seasonal transitions in VR environment.

5.4 Discussion

5.4.1 Limitations of controlled conditions in digital tree modeling

The investigation into digital tree modeling within virtual environments has unveiled significant insights into the modeling process. However, it is important to acknowledge that the controlled conditions used in this study may not fully represent the complexities of natural environments, potentially impacting the applicability of these techniques to real trees. The consistency of backgrounds that benefits the NeRF process might not translate well to the varied backgrounds of real-world settings, potentially affecting model accuracy and effectiveness. For example, in the real park scene shown in the series of images, the quality of the image generated by the stabilized diffusion model deteriorated due to the complexity of the scene, as shown in Figure 5.6, with some noise and broken holes.

Figure 5.6 Example of image quality degradation in complex real-world scenes.

Future research should focus on addressing these limitations and enhancing the robustness of the modeling system. This could involve developing more advanced algorithms capable of handling diverse and complex backgrounds, integrating real-world data on plant growth patterns and local climate conditions, and incorporating sophisticated lighting and atmospheric modeling. Improving the system's ability to accurately simulate seasonal changes in vegetation, including tree height, leaf density, and color variations, will be crucial. Additionally, exploring ways to represent temporal dynamics beyond just seasonal changes, such as daily and weather-related variations, could significantly enhance the realism of the models. The development of adaptive techniques that can adjust to the intricacies of real urban green spaces at various scales, from individual plants to entire ecosystems, will be essential. Furthermore, creating tools that allow urban planners and designers to interact with and modify the generated models in real-time could greatly enhance the practical applicability of this technology in urban planning and design processes. By addressing these challenges, future iterations of the system could provide more accurate, versatile, and practical digital tree models that better represent the complexities of real-world urban green spaces across different seasons and environmental conditions.

5.4.2 *Extending the framework to multi-plant scenes and ecosystems*

Furthermore, the current framework primarily focuses on modeling and rendering individual plants. However, in real-world urban landscapes, plants often exist as communities or ecosystems, exhibiting complex spatial arrangements and interactions. Future work could explore extending the framework to generate multi-plant scenes, considering occlusions, competitions, and mutualistic relationships among plants. This would require the development of more sophisticated scene representation and rendering techniques, as well as the incorporation of knowledge from ecology and plant physiology to guide the model generation process.

5.4.3 *Reducing dependence on extensive real image data*

Additionally, the current approach heavily relies on high-quality plant image data to train the NeRF and Stable diffusion models. However, acquiring large-scale, diverse plant image datasets can be expensive and time-consuming. Future research could investigate leveraging synthetic data or semi-supervised learning techniques to reduce the dependence on extensive real image data, thereby improving the scalability and applicability of the framework. Moreover, integrating this approach with other data sources, such as LiDAR point clouds or multispectral imagery, could provide additional geometric and semantic information to guide the generation of 4D plant models.

5.4.4 *Converting image-space representations into explicit 3D models*

While the plant models generated by the current framework exhibit excellent visual quality and realism, they are still image-based representations lacking explicit 3D structure and topology. To enable more in-depth analysis and interaction, such as immersive exploration in VR or physical simulations, it is necessary to convert the image-space representations into explicit 3D models. This can be achieved by combining advanced 3D reconstruction techniques, such as surface mesh extraction and skeleton estimation, with rule-based procedural modeling approaches.

5.4.5 Integration with urban simulation tools and decision support systems

Lastly, to fully harness the potential of 4D plant models in urban landscape planning and design, integration with other urban simulation tools and decision support systems is crucial. This involves developing user-friendly interfaces and visualization tools that allow urban planners and stakeholders to intuitively explore and analyze different vegetation scenarios and their impacts on the environment and human well-being. Integrating plant growth models with microclimate, hydrology, and biodiversity models can provide a more comprehensive assessment of urban ecosystem services.

5.4.6 Complementary relationship between 4D modeling and S3PVI

It is also essential to consider the complementary relationship between the 4D modeling approach presented in this chapter and the multi-temporal visualization analysis framework introduced in Chapter 4, particularly the S3PVI index. The S3PVI index, which quantifies the visual importance of individual plant species across different seasons, provides a solid theoretical foundation and data support for the pursuit of realism and dynamism in 4D modeling. By revealing the spatiotemporal differentiation patterns of vegetation visual features, S3PVI offers valuable prior knowledge to guide the generation of realistic and diverse vegetation images in the 4D modeling process. Conversely, the 4D modeling approach enhances the multi-temporal visualization analysis by providing a more intuitive and immersive means of expressing the S3PVI results. The synergistic integration of S3PVI and 4D modeling opens up new possibilities for comprehensive understanding and representation of the multi-dimensional spatiotemporal features of urban green spaces. Future research could explore the integration of S3PVI with other ecological and recreational indicators, such as biodiversity, carbon sequestration, and social interaction, to provide a more holistic analysis and visualization of the complex interplay between the visual, ecological, and social functions of urban vegetation within the 4D modeling platform.

5.4.7 Future work and challenges

For future work, the plan is to explore improvements in simulating multiple plants simultaneously, dealing with occlusions and complex structures, improving the realism of seasonal transitions, and introducing MR interactions. By addressing these challenges and integrating the framework with a broader range of urban simulation and planning tools, more accurate, informative, and actionable digital twins of urban green spaces can be created, contributing to the development of sustainable and resilient cities.

Additionally, the research aims to incorporate the S3PVI index introduced in Chapter 4 into the workflow. The process would involve using S3PVI to assess the current situation, followed by planning and designing trees and arranging various plants on a 3D ground plane. The 4D seasonal changes would then be generated using the methods described in this chapter. Finally, the aesthetics of the planned landscape would be evaluated using an enhanced version of S3PVI, incorporating the technological advancements discussed. This integrated approach could provide a more comprehensive and dynamic assessment of urban green spaces across different seasons, potentially leading to more informed decision-making in urban landscape design and management.

5.5 Summary of this chapter

This chapter introduced a framework for 4D plant landscape modeling that integrates time dynamics with advanced imaging techniques. The proposed approach combines NeRF for learning 3D scene representations and stable diffusion for synthesizing time-varying appearances, enabling the generation of realistic plant models across different seasons. The integration of LoRA and ControlNet further enhances the visual quality and geometric consistency of the generated images.

Experiments conducted on a dataset of virtual plant images demonstrated the effectiveness of the proposed framework in capturing both the spatial structure and temporal dynamics of plants. Quantitative evaluation using the PSNR, SSIM, and LPIPS metrics showed that the generated plant models exhibit good fidelity and perceptual similarity compared to the ground truth.

The developed 4D plant modeling framework has potential applications in urban landscape planning and design. By providing realistic visualizations of plant growth and seasonal variations, it can assist urban planners and environmental scientists in making decisions regarding the selection, placement, and management of vegetation in urban spaces. The framework may also be extended to other applications, such as virtual reality simulations, video games, and film production, where realistic plant models are desirable for creating immersive environments.

The 4D modeling approach presented in this chapter and the S3PVI index introduced in Chapter 4 have a complementary relationship. The spatiotemporal differentiation patterns of vegetation visual features revealed by S3PVI can inform the pursuit of realism and dynamism in 4D modeling. The integration of S3PVI data, including multi-temporal vegetation segmentation maps and quantitative descriptions of plant visibility levels, can enhance the accuracy and expressiveness of the generated 4D scenes. Conversely, the 4D modeling method provides a way to visualize and communicate the S3PVI results, potentially facilitating the dissemination of scientific knowledge to stakeholders and the public.

The integration of S3PVI and 4D modeling offers opportunities for understanding and representing the multi-dimensional spatiotemporal features of urban green spaces. It represents a new approach in urban green space research, combining quantitative assessment with realistic visualization. This integrated approach has the potential to support the understanding, planning, and management of urban vegetation, contributing to the development of sustainable cities.

Future work will focus on improving the efficiency and scalability of the proposed framework, exploring alternative architectures for learning 3D representations, and incorporating additional factors such as environmental conditions and plant-plant interactions. Integrating the 4D plant models with urban simulation tools and decision support systems could also be investigated to facilitate urban landscape planning and management. Moreover, future research could explore the integration of S3PVI with other ecological and recreational indicators to provide a more comprehensive analysis and visualization of the interplay between the visual, ecological, and social functions of urban vegetation within the 4D modeling platform.

In conclusion, the proposed 4D plant modeling framework, built upon the foundation of the comprehensive assessment framework and the S3PVI method, offers a tool for visualizing and analyzing the dynamic nature of urban green spaces. The insights gained from these approaches can contribute to a better understanding of urban green space performance, informing planning and management practices. The integration of multi-temporal visualization and 4D modeling represents an advancement in the field of urban landscape modeling, enabling the creation of realistic plant visualizations that capture the temporal dynamics of plant growth and development.

Chapter 6
Conclusion

6.1 Summary

This dissertation presents a comprehensive and innovative approach to assessing and modeling urban green spaces by integrating multi-source data, novel indicators, and advanced visualization technologies. The research aims to address the limitations of existing methods and provide urban planners and policymakers with more comprehensive, nuanced, and actionable insights for evidence-based decision-making in urban green space planning and management.

6.1.1 Comprehensive assessment framework for urban green spaces

To achieve this goal, this study constructed a multidimensional evaluation framework that integrates multiple data sources and evaluation metrics, such as the GVI and GCR, to assess the spatial distribution, visibility, and composition of urban green spaces in a multi-dimensional manner. The case study of Osaka City demonstrates the framework's potential in revealing variations in green space provision and the influence of topography on green space distribution, providing objective and quantitative analysis results as data support for designers and planners.

6.1.2 Multi-temporal urban green space vegetation visualization analysis framework and S3PVI

This study also proposes a multi-temporal urban green space vegetation visualization analysis framework that integrates various data sources and advanced technologies to capture the temporal changes and seasonal variations in vegetation characteristics. The S3PVI indicator aims to evaluate the visual importance and attractiveness of different plants in urban landscapes, providing quantitative basis for plant selection and configuration decisions.

The framework and S3PVI indicator are primarily intended for urban planners, park designers, and street/road planners. In practical applications, these tools can be utilized to first evaluate the current situation with S3PVI, especially in brownfield projects. This step provides objective and detailed baseline data, helping to clarify the project's starting point and areas for improvement. Subsequently, users can plan and design trees, laying out various plants on a 3D ground surface. The framework allows designers to simulate the visual effects of different plant combinations across seasons, thereby optimizing plant selection and spatial arrangement.

The application of this method in Sanshikisaido, Suita City, Osaka Prefecture, Japan, revealed spatiotemporal differentiation patterns of visibility characteristics for different plant types. For instance, by analyzing the seasonal changes in S3PVI for iconic plants such as cherry blossoms and maple trees, planners can optimize plant configurations to ensure year-round landscape attractiveness. This not only demonstrates the practical application value of the method but also provides new perspectives and approaches for urban green space planning and management.

The multi-temporal nature of the framework enables users to predict and visualize vegetation growth and changes over time, which is particularly important for long-term planning. This comprehensive approach to assessment, planning, and management of urban green spaces may support professionals in making more informed and sustainable decisions, potentially leading to the creation of visually appealing and ecologically diverse urban green space systems.

6.1.3 Application of advanced visualization technologies

This study integrates NeRF and Stable Diffusion to create immersive 4D visualizations of urban green spaces, enabling intuitive exploration and analysis of vegetation dynamics in space and time. These technologies aim to enable intuitive exploration and analysis of vegetation dynamics in space and time, potentially offering urban planners and landscape architects a tool for visualizing and communicating design proposals.

The application of these visualization techniques may be particularly useful in the later stages of the design process, after initial assessments and planning have been conducted using tools like S3PVI. Planners and designers could use this 4D modeling approach to create virtual walkthroughs of proposed green spaces, allowing stakeholders to experience how the vegetation might change across seasons and years.

Building upon the data foundation laid by the S3PVI assessment framework, the integration of Low-Rank Adaptation (LoRA) and ControlNet aims to improve the visual quality and geometric consistency of the generated images. This enhancement could potentially allow for more accurate representations of specific plant species and their growth patterns, which may be crucial for detailed design reviews and public presentations.

Experiments conducted on a virtual plant image dataset suggest that this framework might be effective in capturing the spatial structure and temporal dynamics of plants. In practice, this could mean that urban planners and landscape architects might use these tools to simulate and evaluate different planting scenarios over extended periods, potentially informing decisions about species selection, placement, and long-term maintenance strategies.

6.2 Research contributions

6.2.1 Comprehensive assessment framework

The comprehensive assessment framework contributes to the field of urban green space assessment by integrating multi-source data and evaluation metrics to provide a multi-dimensional understanding of urban green spaces. The framework addresses the limitations of existing methods that focus on single dimensions and oversimplify the evaluation process. By considering the interplay between different indicators and the influence of topography on green space distribution, the framework offers a more holistic and context-specific approach to urban green space assessment, supporting evidence-based decision-making in urban planning and management.

6.2.2 Multi-temporal urban green space vegetation visualization analysis framework and S3PVI

The multi-temporal urban green space vegetation visualization analysis framework represents a contribution to the field of urban green space assessment by providing a comprehensive approach to capturing the temporal changes and seasonal variations in vegetation characteristics. The framework integrates various data sources and advanced technologies, enabling the characterization of the spatiotemporal differentiation patterns of vegetation visual features. The introduction of the S3PVI indicator within this framework further enhances its capability to evaluate the visual importance and attractiveness of different plant types in urban landscapes, addressing a research gap in existing assessment methods that focus primarily on overall greenness or vegetation health.

6.2.3 Integration of advanced visualization technologies

This study explores the integration of NeRF and Stable Diffusion to enhance the realism and expressiveness of 4D urban landscape modeling. By combining these visualization techniques, the proposed framework aims to capture both the structural and temporal aspects of urban vegetation with improved detail and consistency.

The incorporation of LoRA and ControlNet into the modeling pipeline is intended to refine the visual quality and spatial coherence of the generated plant images. LoRA's parameter-efficient fine-tuning capabilities are leveraged to utilize semantic information provided by the S3PVI values, aiming to ensure better alignment between the synthesized vegetation and the intended seasonal characteristics. ControlNet's semantic segmentation guidance is employed to maintain geometric consistency by guiding the image manipulation process to respect the underlying plant structures.

The integration of these visualization technologies contributes to the development of tools for exploring and analyzing urban green spaces across spatial and temporal dimensions. The resulting framework has the potential to provide insights into the dynamics of urban ecosystems and support decision-making processes related to urban landscape management and planning.

6.3 Limitations and future work

While this dissertation introduces novel approaches to urban green space assessment and modeling, there are several areas that require further research and development:

1) Integration of the proposed frameworks: Future work should focus on developing an integrated system that combines the comprehensive assessment framework, the multi-temporal visualization analysis framework with S3PVI, and the 4D plant modeling approach. This integration could enable an understanding and representation of urban green spaces, from multi-dimensional assessment to dynamic visualization. For example, a unified software platform could be developed, allowing users to transition seamlessly from initial assessment to detailed planning and long-term simulation. This integrated system could also incorporate machine learning algorithms to optimize data flow and decision support functions between different frameworks.

2) Incorporation of additional data sources: Integrating the proposed frameworks with other data sources, such as LiDAR point clouds, multispectral imagery, and IoT sensor data, could provide a more comprehensive understanding of urban green spaces. This integration could not only assess ecological functions, microclimate effects, and human-environment interactions but also improve model accuracy and predictive capabilities. For instance, incorporating high-resolution remote sensing data could enhance vegetation health assessment, while integrating real-time meteorological data could improve the accuracy of seasonal change predictions. Future research could also explore effective ways to fuse and process these heterogeneous data sources, as well as how to utilize crowdsourced data to supplement traditional data collection methods while preserving privacy.

3) Validation and refinement of the S3PVI: The S3PVI indicator introduced in Chapter 4 should be further validated using a diverse range of urban green space types and cultural contexts. Expanding the current plant dataset to include more species and environmental conditions could improve the robustness and applicability of S3PVI. Collaborating with domain experts and stakeholders could also help refine the indicator and ensure its relevance across different urban landscapes. Future research could consider developing a dynamic S3PVI system capable of automatically

adjusting weights based on local and seasonal factors. Additionally, exploring methods to combine S3PVI with other quantitative and qualitative indicators, such as biodiversity indices, economic metrics, health data, or public preference surveys, might yield a more comprehensive urban green space quality assessment system that addresses the full range of ecological, social, economic, and psychological impacts of urban green spaces.

4) Expansion of the 4D plant modeling framework: The 4D plant modeling approach presented in Chapter 5 could be extended to generate multi-plant scenes, considering factors such as occlusions, competitions, and mutualistic relationships among plants. Incorporating plant growth models and environmental factors could enable more realistic and dynamic visualizations of urban vegetation. Future research directions might include developing more complex ecosystem models capable of simulating long-term dynamics of plant communities and their responses to climate change. Moreover, integrating urban microclimate models and human activity patterns could provide a more comprehensive view of urban ecosystem dynamics. Exploring how to increase model complexity while maintaining computational efficiency will be an important technical challenge.

5) Integration of S3PVI with 4D plant modeling: Future research should focus on developing methods to seamlessly combine these two approaches for a more comprehensive assessment of urban green spaces. Efforts should be directed towards expanding and refining the S3PVI dataset to encompass a broader range of plant species across various seasons and growth stages. Enhancing the S3PVI detection system to accurately identify and assess plants over time will be crucial. Improving the visual quality and accuracy of 4D modeling outputs to meet the high-definition standards required for effective S3PVI analysis is another key area for development. A primary objective will be to develop dynamic assessment capabilities for S3PVI, enabling it to evaluate the temporal dimensions inherent in 4D models. This would allow for a more nuanced understanding of seasonal changes in urban vegetation. Future work could explore innovative ways to integrate S3PVI into the 4D modeling process, potentially using it to guide both model generation and evaluation. Extending this integrated system to larger spatial scales, such as entire cities or regions, while maintaining detail and accuracy, represents another important research direction. This expansion will necessitate the development of efficient algorithms capable of processing large-scale data and performing complex computations. Pursuing these research avenues aims to create a more dynamic and comprehensive approach to urban green space modeling and assessment, ultimately providing urban planners and designers with more sophisticated tools for decision-making and strategy development.

6) Integration with urban planning and management tools: To fully leverage the potential of the proposed frameworks, integration with existing urban planning and management tools is crucial. Developing user-friendly interfaces and visualization

platforms that allow stakeholders to explore and analyze urban green spaces could facilitate evidence-based decision-making and collaborative planning processes. Future research could explore how to seamlessly integrate these new tools into existing urban information systems and decision support frameworks. For example, developing interactive platforms that support multi-stakeholder engagement, or creating dynamic urban green space "digital twin" systems capable of real-time updates. Additionally, investigating how to utilize MR and virtual VR technologies to enhance stakeholder engagement and public consultation processes could also be a valuable direction.

Addressing these limitations and exploring the identified future research directions has the potential to significantly advance the field of urban green space assessment and modeling. The development of a more comprehensive, data-driven, and stakeholder-oriented approach could bridge the gaps between assessment, visualization, and planning, ultimately supporting the creation of sustainable, resilient, and livable urban environments. This integrated approach could enhance decision support by providing urban planners and policymakers with more nuanced and actionable insights, leading to more informed decision-making processes. Moreover, the integration of advanced visualization technologies with user-friendly interfaces could facilitate better communication and collaboration among diverse stakeholders, fostering more inclusive and participatory urban planning processes.

The incorporation of temporal dynamics and ecological relationships in the modeling frameworks may enable better long-term planning and management of urban green spaces, enhancing their resilience to environmental changes and societal needs. Additionally, the proposed integrated system could support adaptive management strategies by providing real-time data and predictive modeling capabilities, allowing for more responsive and flexible urban green space management.

This dissertation presents a comprehensive and innovative approach to urban green space assessment and modeling, integrating multi-source data, novel indicators, and advanced visualization technologies. The proposed frameworks and methodologies contribute to the development of more holistic, nuanced, and actionable approaches to understanding and managing urban green spaces. These contributions include a comprehensive assessment framework that integrates multiple data sources and evaluation metrics, a multi-temporal urban green space vegetation visualization analysis framework featuring the novel S3PVI indicator, and the integration of advanced visualization technologies for dynamic 4D modeling of urban green spaces.

Collectively, these contributions support the creation of sustainable and resilient urban environments by providing tools for more informed decision-making and adaptive management. However, realizing the full potential of these approaches requires addressing the identified limitations and pursuing the outlined future research directions. This necessitates continued research efforts and collaboration among diverse stakeholders, including urban planners, environmental scientists, data scientists, and local communities.

Future work should focus on developing an integrated system that seamlessly combines these frameworks, incorporates a wider range of data sources, and interfaces effectively with existing urban planning and management tools. While the proposed frameworks and methodologies offer potential tools for urban planning and management, their practical impact and effectiveness require further validation and refinement through real-world applications and collaborations with urban planners, policymakers, and local communities. The true value of these approaches will ultimately be determined by their ability to address the complex challenges faced by cities in creating and maintaining sustainable, resilient, and livable green spaces. As urban areas continue to evolve and face various environmental challenges, continued research and interdisciplinary collaboration remain essential to develop and improve methods for urban green space assessment and management.

References

Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., & Weaver, J. (2010). Google Street View: Capturing the World at Street Level. *Computer*, *43*(6), 32–38. https://doi.org/10.1109/MC.2010.170

Aoki, Y., Yasuoka, Y., & Naito, M. (1985). Assessing the impression of street-side greenery. *Landscape Research*, *10*(1), 9–13. https://doi.org/10.1080/01426398508706131

Baidu Map open platform. (n.d.). Retrieved June 2, 2024, from https://lbsyun.baidu.com/

Balram, S., & Dragievi, S. (2005). Attitudes toward urban green spaces: Integrating questionnaire survey and collaborative GIS techniques to improve attitude measurements. *Landscape and Urban Planning*, *71*(2–4), 147–162. https://doi.org/10.1016/S0169-2046(04)00052-0

Basic Plan of Greening in Osaka City. (2013, November). Osaka City. https://www.city.osaka.lg.jp/kensetsu/page/0000239835.html

Beer, A. R., Delshammar, T., & Schildwacht, P. (2003). A Changing Understanding of the Role of Greenspace in High-density Housing: A European Perspective. *Built Environment*, *29*(2), 132–143. https://doi.org/10.2148/benv.29.2.132.54468

Bielinis, E., Takayama, N., Boiko, S., Omelan, A., & Bielinis, L. (2018). The effect of winter forest bathing on psychological relaxation of young Polish adults. *Urban Forestry & Urban Greening*, *29*, 276–283. https://doi.org/10.1016/j.ufug.2017.12.006

Bmaltais. (n.d.). *Kohya's GUI* [Python]. Retrieved September 1, 2023, from https://github.com/bmaltais/kohya_ss (Original work published 2022)

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, *62*(3), 241–252. https://doi.org/10.1016/S0034-4257(97)00104-1

Carmen, R., Jacobs, S., Leone, M., Palliwoda, J., Pinto, L., Misiune, I., Priess, J. A., Pereira, P., Wanner, S., Ferreira, C. S., & Ferreira, A. (2020). Keep it real: Selecting realistic sets of urban green space indicators. *Environmental Research Letters*, *15*(9), 095001. https://doi.org/10.1088/1748-9326/ab9465

Carrus, G., Scopelliti, M., Lafortezza, R., Colangelo, G., Ferrini, F., Salbitano, F., Agrimi, M., Portoghesi, L., Semenzato, P., & Sanesi, G. (2015). Go greener, feel better? The positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landscape and Urban Planning*, *134*, 221–228. https://doi.org/10.1016/j.landurbplan.2014.10.022

Chavan, G. (2023). *Identification of Plant Species using Remote Sensing Techniques: A Review*. *29*(9).

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). *Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation* (arXiv:1802.02611). arXiv. http://arxiv.org/abs/1802.02611

Cilliers, J., & Cilliers, S. (2015). From green to gold: A South African example of valuing urban green spaces in some residential areas in Potchefstroom. *Town and Regional Planning*, *67*.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). *The Cityscapes Dataset for Semantic Urban Scene Understanding* (arXiv:1604.01685). arXiv. http://arxiv.org/abs/1604.01685

Daniels, B., Zaunbrecher, B. S., Paas, B., Ottermanns, R., Ziefle, M., & Roß-Nickoll, M. (2018). Assessment of urban green space structures and their quality from a multidimensional perspective. *Science of The Total Environment*, *615*, 1364–1378. https://doi.org/10.1016/j.scitotenv.2017.09.167

Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Du, H., Jiang, H., Song, X., Zhan, D., & Bao, Z. (2016). Assessing the Visual Aesthetic Quality of Vegetation Landscape in Urban Green Space from a Visitor's Perspective. *Journal of Urban Planning and Development*, *142*(3), 04016007. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000329

Dutta, D., Rahman, A., Paul, S. K., & Kundu, A. (2022). Spatial and temporal trends of urban green spaces: An assessment using hyper-temporal NDVI datasets. *Geocarto International*, *37*(25), 7983–8003. https://doi.org/10.1080/10106049.2021.1989499

Elmqvist, T., Setälä, H., Handel, S., Van Der Ploeg, S., Aronson, J., Blignaut, J., Gómez-Baggethun, E., Nowak, D., Kronenberg, J., & De Groot, R. (2015). Benefits of restoring ecosystem services in urban areas. *Current Opinion in Environmental Sustainability*, *14*, 101–108. https://doi.org/10.1016/j.cosust.2015.05.001

FFmpeg Developers. (2016). http://ffmpeg.org/

Fix, P. J., Brooks, J. J., & Harrington, A. M. (2018). Achieving goals and making meanings: Toward a unified model of recreational experience. *Journal of Outdoor Recreation and Tourism*, *23*, 16–25. https://doi.org/10.1016/j.jort.2018.04.004

Gomes, D. P. S., & Zheng, L. (2020). *Leaf Segmentation and Counting with Deep Learning: On Model*

Certainty, Test-Time Augmentation, Trade-Offs (arXiv:2012.11486). arXiv. http://arxiv.org/abs/2012.11486

Google Earth. (n.d.). Retrieved September 1, 2023, from https://earth.google.com/web/@37.40790556,140.50031482,383.11662861a,0d,41.53484368y,-139.84091982h,95.73772503t,0r

Grunewald, K., & Bastian, O. (2015). *Ecosystem Services – Concept, Methods and Case Studies*. Springer.

Gupta, K., Kumar, P., Pathan, S. K., & Sharma, K. P. (2012). Urban Neighborhood Green Index – A measure of green spaces in urban areas. *Landscape and Urban Planning*, *105*(3), 325–335. https://doi.org/10.1016/j.landurbplan.2012.01.003

Haaland, C., & Van Den Bosch, C. K. (2015). Challenges and strategies for urban green-space planning in cities undergoing densification: A review. *Urban Forestry & Urban Greening*, *14*(4), 760–771. https://doi.org/10.1016/j.ufug.2015.07.009

Han, Y., Zhong, T., Yeh, A. G. O., Zhong, X., Chen, M., & Lü, G. (2023). Mapping seasonal changes of street greenery using multi-temporal street-view images. *Sustainable Cities and Society*, *92*, 104498. https://doi.org/10.1016/j.scs.2023.104498

Hartig, T., Mitchell, R., De Vries, S., & Frumkin, H. (2014). Nature and Health. *Annual Review of Public Health*, *35*(1), 207–228. https://doi.org/10.1146/annurev-publhealth-032013-182443

Hore, A., & Ziou, D. (2010). Image Quality Metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition*, 2366–2369. https://doi.org/10.1109/ICPR.2010.579

Hoyle, H., Hitchmough, J., & Jorgensen, A. (2017). All about the 'wow factor'? The relationships between aesthetics, restorative effect and perceived biodiversity in designed urban planting. *Landscape and Urban Planning*, *164*, 109–123. https://doi.org/10.1016/j.landurbplan.2017.03.011

Hu, A., Yabuki, N., & Fukuda, T. (2023). *Development of a Method for Assessing the View Index of Plants of Interest Using Deep Learning*. 585–594. https://doi.org/10.52842/conf.caadria.2023.1.585

Hu, A., Yabuki, N., Fukuda, T., Kaga, H., Takeda, S., & Matsuo, K. (2023). Harnessing multiple data sources and emerging technologies for comprehensive urban green space evaluation. *Cities*, *143*, 104562. https://doi.org/10.1016/j.cities.2023.104562

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2021). *LoRA: Low-Rank Adaptation of Large Language Models* (arXiv:2106.09685). arXiv. http://arxiv.org/abs/2106.09685

Huang, C., Yang, J., Lu, H., Huang, H., & Yu, L. (2017). Green Spaces as an Indicator of Urban Health: Evaluating Its Changes in 28 Mega-Cities. *Remote Sensing*, *9*(12), Article 12. https://doi.org/10.3390/rs9121266

Huang, L., Yuan, Y., Guo, J., Zhang, C., Chen, X., & Wang, J. (2019). *Interlaced Sparse Self-Attention for Semantic Segmentation* (arXiv:1907.12273). arXiv. http://arxiv.org/abs/1907.12273

James, P., Tzoulas, K., Adams, M. D., Barber, A., Box, J., Breuste, J., Elmqvist, T., Frith, M., Gordon, C., Greening, K. L., Handley, J., Haworth, S., Kazmierczak, A. E., Johnston, M., Korpela, K., Moretti, M., Niemelä, J., Pauleit, S., Roe, M. H., … Ward Thompson, C. (2009). Towards an integrated understanding of green space in the European built environment. *Urban Forestry & Urban Greening*, *8*(2), 65–75. https://doi.org/10.1016/j.ufug.2009.02.001

Kameoka, T., Uchida, A., Sasaki, Y., & Ise, T. (2022). Assessing streetscape greenery with deep neural network using Google Street View. *Breeding Science*, *72*(1), 107–114. https://doi.org/10.1270/jsbbs.21073

Kandt, J., & Batty, M. (2021). Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities*, *109*, 102992. https://doi.org/10.1016/j.cities.2020.102992

Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). *3D Gaussian Splatting for Real-Time Radiance Field Rendering* (arXiv:2308.04079). arXiv. http://arxiv.org/abs/2308.04079

Kingma, D. P., & Ba, J. (2017). *Adam: A Method for Stochastic Optimization* (arXiv:1412.6980). arXiv. http://arxiv.org/abs/1412.6980

Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., & Houlsby, N. (2020). Big Transfer (BiT): General Visual Representation Learning. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Computer Vision – ECCV 2020* (pp. 491–507). Springer International Publishing. https://doi.org/10.1007/978-3-030-58558-7_29

Kotowska, D., Pärt, T., & Żmihorski, M. (2021). Evaluating Google Street View for tracking invasive alien plants along roads. *Ecological Indicators*, *121*, 107020. https://doi.org/10.1016/j.ecolind.2020.107020

Kuper, R. (2015). Preference, Complexity, and Color Information Entropy Values for Visual Depictions of Plant and Vegetative Growth. *HortTechnology*, *25*(5), 625–634. https://doi.org/10.21273/HORTTECH.25.5.625

Li, X. (2021). Examining the spatial distribution and temporal change of the green view index in New York City using Google Street View images and deep learning. *Environment and Planning B: Urban Analytics and City Science*, *48*(7), 2039–2054. https://doi.org/10.1177/2399808320962511

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, *14*(3), 675–685. https://doi.org/10.1016/j.ufug.2015.06.006

Liang, X., Zhao, T., & Biljecki, F. (2023). Revealing spatio-temporal evolution of urban visual environments with street view imagery. *Landscape and Urban Planning*, *237*, 104802. https://doi.org/10.1016/j.landurbplan.2023.104802

Lindemann-Matthies, P., & Brieger, H. (2016). Does urban gardening increase aesthetic quality of urban areas? A case study from Germany. *Urban Forestry & Urban Greening*, *17*, 33–41. https://doi.org/10.1016/j.ufug.2016.03.010

List of Distinctive Town Names in Osaka City. (2022, December 27). Osaka City. https://www.city.osaka.lg.jp/shimin/page/0000061040.html

Lovell, S. T., & Taylor, J. R. (2013). Supplying urban ecosystem services through multifunctional green infrastructure in the United States. *Landscape Ecology*, *28*(8), 1447–1463. https://doi.org/10.1007/s10980-013-9912-y

Lu, F., Xu, Y., Chen, G., Li, H., Lin, K.-Y., & Jiang, C. (2023). *Urban Radiance Field Representation with Deformable Neural Mesh Primitives* (arXiv:2307.10776). arXiv. http://arxiv.org/abs/2307.10776

Luederitz, C., Brink, E., Gralla, F., Hermelingmeier, V., Meyer, M., Niven, L., Panzer, L., Partelow, S., Rau, A.-L., Sasaki, R., Abson, D. J., Lang, D. J., Wamsler, C., & von Wehrden, H. (2015). A review of urban ecosystem services: Six key challenges for future research. *Ecosystem Services*, *14*, 98–112. https://doi.org/10.1016/j.ecoser.2015.05.001

Maclean, I. M. D., Duffy, J. P., Haesen, S., Govaert, S., De Frenne, P., Vanneste, T., Lenoir, J., Lembrechts, J. J., Rhodes, M. W., & Van Meerbeek, K. (2021). On the measurement of microclimate. *Methods in Ecology and Evolution*, *12*(8), 1397–1410. https://doi.org/10.1111/2041-210X.13627

Ministry of Land, Infrastructure, Transport and Tourism. (2008). Geospatial Information Authority of Japan. https://www.gsi.go.jp/

Mouragnon, E., Lhuillier, M., Dhome, M., Dekeyser, F., & Sayd, P. (2006). Real Time Localization and 3D Reconstruction. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 363–370. https://doi.org/10.1109/CVPR.2006.236

Nitoslawski, S. A., Galle, N. J., Van Den Bosch, C. K., & Steenberg, J. W. N. (2019). Smarter ecosystems for smarter cities? A review of trends, technologies, and turning points for smart urban forestry. *Sustainable Cities and Society*, *51*, 101770. https://doi.org/10.1016/j.scs.2019.101770

Noland, R. B., Weiner, M. D., Gao, D., Cook, M. P., & Nelessen, A. (2017). Eye-tracking technology, visual preference surveys, and urban design: Preliminary evidence of an effective methodology. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *10*(1), 98–110. https://doi.org/10.1080/17549175.2016.1187197

Oh, C., Jang, Y., Shim, D., Kim, C., Kim, J., & Kim, H. J. (2024). Automatic Pseudo-LiDAR Annotation: Generation of Training Data for 3D Object Detection Networks. *IEEE Access*, *12*, 14227–14237. https://doi.org/10.1109/ACCESS.2024.3355137

Osaka City: 2020 Population Census Results. (2020). https://www.city.osaka.lg.jp/shisei/category/3055-2-3-2-6-0-0-0-0-0.html

Osaka Prefectural Government. (2020). Guidelines for Green Vision Survey. https://www.pref.osaka.lg.jp/kannosomu/ryokushiritsu/

Pauleit, S., Ambrose-Oji, B., Andersson, E., Anton, B., Buijs, A., Haase, D., Elands, B., Hansen, R., Kowarik, I., Kronenberg, J., Mattijssen, T., Stahl Olafsson, A., Rall, E., Van Der Jagt, A. P. N., & Konijnendijk Van Den Bosch, C. (2019). Advancing urban green infrastructure in Europe: Outcomes and reflections from the GREEN SURGE project. *Urban Forestry & Urban Greening*, *40*, 4–16. https://doi.org/10.1016/j.ufug.2018.10.006

Pulighe, G., Fava, F., & Lupia, F. (2016). Insights and opportunities from mapping ecosystem services of urban green spaces and potentials in planning. *Ecosystem Services*, *22*, 1–10.

https://doi.org/10.1016/j.ecoser.2016.09.004

Riechers, M., Strack, M., Barkmann, J., & Tscharntke, T. (2019). Cultural Ecosystem Services Provided by Urban Green Change along an Urban-Periurban Gradient. *Sustainability*, *11*(3), 645. https://doi.org/10.3390/su11030645

Schonberger, J. L., & Frahm, J.-M. (2016). Structure-from-Motion Revisited. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4104–4113. https://doi.org/10.1109/CVPR.2016.445

Seiferling, I., Naik, N., Ratti, C., & Proulx, R. (2017). Green streets − Quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning*, *165*, 93–101. https://doi.org/10.1016/j.landurbplan.2017.05.010

Shao, Z., Sumari, N. S., Portnov, A., Ujoh, F., Musakwa, W., & Mandela, P. J. (2021). Urban sprawl and its impact on sustainable urban development: A combination of remote sensing and social media data. *Geo-Spatial Information Science*, *24*(2), 241–255. https://doi.org/10.1080/10095020.2020.1787800

Shiraishi, K., & Terada, T. (2024). Tokyo's urban tree challenge: Decline in tree canopy cover in Tokyo from 2013 to 2022. *Urban Forestry & Urban Greening*, *97*, 128331. https://doi.org/10.1016/j.ufug.2024.128331

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*(1), 60. https://doi.org/10.1186/s40537-019-0197-0

Sicard, P., Coulibaly, F., Lameiro, M., Araminiene, V., De Marco, A., Sorrentino, B., Anav, A., Manzini, J., Hoshika, Y., Moura, B. B., & Paoletti, E. (2023). Object-based classification of urban plant species from very high-resolution satellite imagery. *Urban Forestry & Urban Greening*, *81*, 127866. https://doi.org/10.1016/j.ufug.2023.127866

Sodjinou, S. G., Mohammadi, V., Sanda Mahama, A. T., & Gouton, P. (2022). A deep semantic segmentation-based algorithm to segment crops and weeds in agronomic color images. *Information Processing in Agriculture*, *9*(3), 355–364. https://doi.org/10.1016/j.inpa.2021.08.003

Song, Y., Huang, B., Cai, J., & Chen, B. (2018). Dynamic assessments of population exposure to urban greenspace using multi-source big data. *Science of The Total Environment*, *634*, 1315–1325. https://doi.org/10.1016/j.scitotenv.2018.04.061

Stessens, P., Canters, F., Huysmans, M., & Khan, A. Z. (2020). Urban green space qualities: An integrated approach towards GIS-based assessment reflecting user perception. *Land Use Policy*, *91*, 104319. https://doi.org/10.1016/j.landusepol.2019.104319

Tan, M., & Le, Q. V. (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks* (arXiv:1905.11946). arXiv. http://arxiv.org/abs/1905.11946

Ulrich, R. S. (1986). Human responses to vegetation and landscapes. *Landscape and Urban Planning*, *13*, 29–44. https://doi.org/10.1016/0169-2046(86)90005-8

Urban area of Osaka City. (1955). Osaka City. https://www.city.osaka.lg.jp/seisakukikakushitsu/page/0000010265.html

Wang, R., Zhao, J., & Liu, Z. (2016). Consensus in visual preferences: The effects of aesthetic quality and landscape types. *Urban Forestry & Urban Greening*, *20*, 210–217. https://doi.org/10.1016/j.ufug.2016.09.005

Wang, X., Rodiek, S., Wu, C., Chen, Y., & Li, Y. (2016). Stress recovery and restorative effects of viewing different urban park scenes in Shanghai, China. *Urban Forestry & Urban Greening*, *15*, 112–122. https://doi.org/10.1016/j.ufug.2015.12.003

Weier, J., & Herring, D. (2000). *Measuring Vegetation (NDVI & EVI)* [Text.Article]. NASA Earth Observatory. https://earthobservatory.nasa.gov/features/MeasuringVegetation

Wellmann, T., Schug, F., Haase, D., Pflugmacher, D., & Van Der Linden, S. (2020). Green growth? On the relation between population density, land use and vegetation cover fractions in a city using a 30-years Landsat time series. *Landscape and Urban Planning*, *202*, 103857. https://doi.org/10.1016/j.landurbplan.2020.103857

Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough.' *Landscape and Urban Planning*, *125*, 234–244. https://doi.org/10.1016/j.landurbplan.2014.01.017

Xia, Y., Yabuki, N., & Fukuda, T. (2021). Development of a system for assessing the quality of urban street-level greenery using street view images and deep learning. *Urban Forestry & Urban Greening*, *59*, 126995. https://doi.org/10.1016/j.ufug.2021.126995

Xiao, Y., Wang, Z., Li, Z., & Tang, Z. (2017). An assessment of urban park access in Shanghai – Implications for the social equity in urban China. *Landscape and Urban Planning*, *157*, 383–393. https://doi.org/10.1016/j.landurbplan.2016.08.007

Xu, Z., Zhou, Y., Wang, S., Wang, L., Li, F., Wang, S., & Wang, Z. (2020). A Novel Intelligent Classification Method for Urban Green Space Based on High-Resolution Remote Sensing Images. *Remote Sensing*, *12*(22), Article 22. https://doi.org/10.3390/rs12223845

Xue, H., Liu, C., Wan, F., Jiao, J., Ji, X., & Ye, Q. (2019). DANet: Divergent Activation for Weakly Supervised Object Localization. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 6588–6597. https://doi.org/10.1109/ICCV.2019.00669

Yao, M., Huo, Y., Ran, Y., Tian, Q., Wang, R., & Wang, H. (2024). *Neural Radiance Field-based Visual Rendering: A Comprehensive Review* (arXiv:2404.00714). arXiv. http://arxiv.org/abs/2404.00714

Ye, Y., Richards, D., Lu, Y., Song, X., Zhuang, Y., Zeng, W., & Zhong, T. (2019). Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landscape and Urban Planning*, *191*, 103434. https://doi.org/10.1016/j.landurbplan.2018.08.028

Zhang, J., & Hu, A. (2022). Analyzing green view index and green view index best path using Google street view and deep learning. *Journal of Computational Design and Engineering*, *9*(5), 2010–2023. https://doi.org/10.1093/jcde/qwac102

Zhang, L., Rao, A., & Agrawala, M. (2023). *Adding Conditional Control to Text-to-Image Diffusion Models* (arXiv:2302.05543). arXiv. http://arxiv.org/abs/2302.05543

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 586–595. https://doi.org/10.1109/CVPR.2018.00068

Zhang, X., Du, S., & Wang, Q. (2017). Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *132*, 170–184. https://doi.org/10.1016/j.isprsjprs.2017.09.007

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). *Pyramid Scene Parsing Network* (arXiv:1612.01105). arXiv. http://arxiv.org/abs/1612.01105

Zhou, D., Zhao, S., Zhang, L., & Liu, S. (2016). Remotely sensed assessment of urbanization effects on vegetation phenology in China's 32 major cities. *Remote Sensing of Environment*, *176*, 272–281. https://doi.org/10.1016/j.rse.2016.02.010

Zhu, H., Nan, X., Yang, F., & Bao, Z. (2023). Utilizing the green view index to improve the urban street greenery index system: A statistical study using road patterns and vegetation structures as entry points. *Landscape and Urban Planning*, *237*, 104780. https://doi.org/10.1016/j.landurbplan.2023.104780