Research on Collaborative Strategies for Cultural Heritage and Sustainable Renewal in Old Urban Areas Based on Multimodal Artificial Intelligence

Boyang Zhang*

Edinburgh College of Art, Edinburgh, United Kingdom

*zhangboyang_1996@163.com

Abstract. This study addresses the challenge of balancing the protection and sustainable renewal of cultural heritage in historical urban areas and constructs a multimodal AI system integrating image recognition, semantic analysis, and GIS data. Through an empirical study of the Pingjiang Historical District in Suzhou, it was found that the core protection area (accounting for 22% of the area) carries 68.5% of the cultural value. Although the temperature in the traditional building-dense area is 2.8℃ higher in summer than that in the modern district, the energy consumption per unit area is 18% lower, demonstrating the value of traditional ecological wisdom. The research established a dual-objective optimization model of cultural value and sustainable benefits, and generated three types of differentiated update schemes through multi-objective genetic algorithms. The conservative update plan achieved a cultural value retention rate of 96% and a carbon reduction rate of 18%. The moderate update plan reached a retention rate of 90% and a carbon reduction rate of 32%. The comprehensive update plan achieved a retention rate of 79% and a carbon reduction rate of 48%. Pareto Frontier analysis indicates that there exists an optimal equilibrium point within the range of 85% to 95% for cultural value retention rates. This study is the first to combine cultural semantic understanding with climate adaptability analysis, breaking through the qualitative limitations of traditional conservation planning. This study established a quantifiable collaborative decision-making framework, providing data support for the renewal of historical urban areas. It confirmed that through precise differentiated strategies, the synergy between cultural heritage protection and green renewal can be achieved. This method can provide reference for historical urban areas around the world that are facing similar challenges.

Keywords: Multimodal artificial intelligence, cultural heritage protection, sustainable renewal, historical urban areas, climate adaptability, green transformation

T. Chamaratana et al. (eds.), *Proceedings of the 2026 12th International Conference on Digital Humanities and Frontiers in Social Sciences (DHFSS 2026)*, Advances in Social Science, Education and Humanities Research 1031,
https://doi.org/10.2991/978-2-38476-611-6_92

1 Research Background and Problem Presentation

As spatial carriers of cultural memory, historical urban areas are facing unprecedented pressure for protection and development [1]. According to the report of UNESCO, over 300 historical cities around the world are under pressure for renewal, among which approximately 40% have experienced the loss of cultural characteristics during the renewal process. Under the backdrop of China's new urbanization, this contradiction is particularly prominent [2]. Data from the Ministry of Housing and Urban-Rural Development shows that over 1,200 historical urban areas across the country need to be updated, among which approximately 60% are simultaneously confronted with the dual challenges of cultural heritage protection and climate adaptability improvement.

1.1 The predicaments Faced by Cultural Heritage Protection

At present, there are mainly three problems in the protection of historical urban areas. Firstly, traditional protection methods overly rely on qualitative evaluation, and the quantitative assessment of cultural value is insufficient, resulting in strong subjectivity in protection decisions [3]. Take the Pingjiang Historical District in Suzhou as an example. Apart from 68 cultural relics protection units, a large number of traditional residences with potential value are facing the threat of renovation due to the lack of scientific assessment standards. Secondly, the demarcation of protection areas is often based on administrative boundaries rather than the distribution of cultural values, resulting in a spatial mismatch between protection and development. Thirdly, the static protection model is difficult to adapt to the demands of dynamic development, leading to an opposition between protection and improvement of people's livelihood.

1.2 Deviations in Sustainable Updates

In terms of sustainable development, the renewal of historical urban areas shows a distinct technicalist tendency [4]. On the one hand, it simply applies the green standards of newly built areas, ignoring the particularity of historical districts. For instance, some renovation projects blindly adopt modern insulation materials, which instead undermine the breathing performance of traditional buildings. On the other hand, the renewal measures lack cultural sensitivity. For instance, the excessive pursuit of energy-saving targets has led to damage to the historical appearance. According to our initial investigation, approximately 35% of the "green renovation" projects have had an irreversible impact on the authenticity of historical buildings.

1.3 New Challenges in Climate Adaptation

Climate change has brought new threats to historical urban areas [5]. Monitoring data from Pingjiang District shows that the frequency of extreme summer heat events has increased by 47% in the past decade, and the risk of urban flooding has risen by 32%. However, the current protection plan lacks targeted climate adaptation strategies.

Meanwhile, the climate adaptability wisdom of traditional buildings (such as tunnel ventilation, water body regulation, etc.) is gradually being lost in the process of modernization, further exacerbating their vulnerability.

1.4 Clear presentation of Research Questions

Based on the above background, this study focuses on the core issue: How to establish a collaborative mechanism for the protection and sustainable renewal of cultural heritage through technological innovation? It is specifically decomposed into three sub-problems:

How to quantitatively evaluate the multi-dimensional cultural value of historical urban areas, especially the implicit cultural semantics?

How to build a multi-objective optimization model that takes into account both cultural protection and climate adaptability?

How to generate spatially differentiated update strategies to achieve a precise balance between protection and development?

This research aims to break through the technical bottlenecks of traditional planning through multimodal AI methods and provide new solutions for the sustainable development of historical urban areas. The Pingjiang Historical District in Suzhou was chosen as a case not only because of its typicality, but also because its complete archival materials and monitoring data provide a good foundation for method verification. The research results have reference value for the protection and renewal of similar historical urban areas around the world.

2 Framework Design of Multimodal AI Systems

2.1 System Architecture

The multimodal artificial intelligence system constructed in this study consists of three closely collaborating core modules, forming a complete decision support system for cultural heritage assessment and renewal.

In terms of the spatial form recognition module, this module is based on deep learning image recognition technology, including CNN convolutional neural networks and Transformer architectures. By analyzing multi-source image data such as satellite images, street view pictures, and building facade photos, it automatically identifies key features such as historical architectural styles, street scales, spatial layouts, and building materials. Empirical research shows that the building age classification model based on the ResNet-152 architecture achieves an accuracy rate of 92.3%, providing a reliable quantitative basis for the value assessment of historical buildings.

The cultural semantic understanding module employs natural language processing technology, especially based on the BERT series of pre-trained models, to conduct in-depth mining of unstructured text data such as historical documents, local Chronicles, residents' oral histories, and social media comments. This module is capable of parsing the profound meanings of cultural symbols, analyzing community emotional tendencies, and understanding the cultural expressions of place spirit. In the performance test, this module achieved an outstanding performance with an F1 score of 0.87 in the task of cultural value sentiment analysis.

The urban GIS and climate data analysis module couples geographic information systems with climate models, integrating multi-dimensional information such as three-dimensional urban models, meteorological data, energy consumption records, and carbon emission data. It focuses on analyzing key indicators such as the heat island effect, energy consumption efficiency, rainwater management, and ecological footprint. This module focuses on the impact of climate change on the protection of cultural heritage and provides a scientific basis for green renewal strategies.

2.2 Multimodal Data Fusion Mechanism

The system adopts an advanced cross-modal attention mechanism to achieve efficient alignment and deep complementarity of three types of heterogeneous data. In terms of spatial-semantic alignment, by establishing the mapping relationship between the physical form of buildings and their cultural meanings, the system can understand the cultural values carried by specific architectural features. In the environment-cultural correlation analysis, the system explores the intrinsic connection between climate data and the vulnerability of cultural Spaces, and identifies which cultural heritages are more vulnerable to the threat of climate change. Meanwhile, the system has also achieved the integration of the time dimension. By analyzing the laws of historical evolution, it predicts future development trends and provides support for the establishment of a long-term protection mechanism. This multimodal fusion mechanism ensures that the system can comprehensively and dynamically assess the value of cultural heritage and provide a data basis for balanced decision-making between protection and renewal.

3 Methodology

The dual-objective optimization model is the core methodology of this study, aiming to achieve the coordinated optimization of the two major goals of cultural heritage protection and green renewal through quantitative analysis. This model first constructs a cultural protection value assessment system, which includes three first-level indicators of historical value, artistic value and social value, as well as 12 second-level indicators under them. Historical value is mainly measured by the age of the building, the correlation with historical events and the relevance to famous people. Artistic value focuses on the representativeness of architectural style, the uniqueness of craftsmanship and the integrity of aesthetics. Social value, on the other hand, pays attention to community recognition, functional continuity and emotional attachment. The weights of each indicator are determined through the Analytic Hierarchy Process (AHP), and quantitatively evaluated using multi-source data (such as literature analysis, image recognition, and semantic analysis).

Meanwhile, the model has established a sustainable renewal benefit assessment system, with a focus on incorporating climate adaptability indicators, including specific dimensions such as energy efficiency, carbon reduction, degree of thermal stress relief, reduction of waterlogging risk, and ecological benefits. For instance, it is required that the energy consumption of buildings be reduced by no less than 25%, the carbon emissions per unit area be decreased by more than 30%, and the proportion of green space be increased by more than 15%, so as to reflect the substantive benefits of green renewal.

On the dual-objective optimization algorithm, the multi-objective genetic algorithm (NSGA-II) is adopted to solve the Pareto optimal solution set. The algorithm takes update intensity, material selection and the degree of technology application as decision variables, and the threshold of cultural value loss and the upper limit of carbon emissions as constraints, simultaneously maximizing the retention rate of cultural value and sustainable benefits. Through iterative optimization, a series of non-dominated solutions are generated, forming a gradient strategy from conservative update to comprehensive update, providing differentiated solutions for different regions. This model not only resolves the trade-off between cultural protection and green renewal, but also transforms traditional qualitative experience into quantifiable scientific decision-making basis, significantly enhancing the systematicness and adaptability of the renewal strategy for historical urban areas.

4 Case Application and Data Analysis

4.1 Research Area Selection and Overview

This study selects the Pingjiang Historical District in Suzhou City, China as the empirical research object. Pingjiang District is the most well-preserved and largest historical district within the ancient city of Suzhou, covering a total area of approximately 1.16 square kilometers, with a core protection area of 0.48 square kilometers. This area was first built in the Southern Song Dynasty and still maintains a double-chessboard layout of "parallel water and land, adjacent rivers and streets". It has one World Cultural Heritage garden, three key national cultural relics protection units, 45 controlled and protected buildings, 18 ancient Bridges, and a well-preserved group of traditional Ming and Qing Dynasty residential buildings. On the one hand, its rich historical and cultural layers (Song and Yuan foundation sites, Ming and Qing architecture, and the style of the Republic of China) provide multi-dimensional samples for cultural value assessment. On the other hand, as a living heritage area that still retains its residential function, it is confronted with sustainable development pressures such as aging infrastructure, high population density (about 23,000 people per square kilometer), prominent summer heat island effect (3-4°C higher than that of new towns), and the risk of waterlogging, perfectly embodying the contradiction between protection and renewal[6].

4.2 Multimodal Data acquisition and Processing flow

This study constructed a dataset covering three dimensions: spatial, semantic, and environmental. In terms of spatial data, a three-dimensional model of the real scene with a resolution of 0.05 meters was obtained through oblique photography by drones. 8,600 street view images and 5,200 building facade photos were collected, and the 1:500 topographic map of Suzhou City and the building age information were integrated. In terms of semantic data, the system has collated 470,000 words of digital texts of inscriptions and local historical materials [7], collected audio of oral history narrated in the Suzhou dialect, and obtained 120,000 tourist comments with geographical tags and 80,000 words of resident interview records through social media apis. The environmental data includes hourly data (temperature, humidity, wind speed, precipitation) from five meteorological stations in the block and its surrounding areas over the past ten years, energy consumption monitoring data of 164 sample buildings, as well as high-resolution remote sensing inversion of surface temperature, vegetation index, etc.

In data processing, the spatial module adopts the improved Mask R-CNN model to achieve pixel-level segmentation (mIoU up to 0.81) of building facade elements (window lattices, horse-head walls, materials), and identifies the architectural style era sequence based on the attention mechanism. The semantic module uses the domain-pre-trained BERT-Arch model to construct a "Cultural value dictionary", quantitatively analyzing implicit dimensions such as place attachment and historical narrative in the text. The environmental module generates a 5-meter grid-level climate map (temperature, wind environment, and sunlight) through GIS spatial interpolation. Ultimately, through feature-level fusion, the three types of data were unified to the same geographic coordinate system to form a multimodal database covering 2,356 basic analysis units (building plots) [8].

4.3 Quantitative Analysis Results of Cultural Values and Climate Vulnerability

Through multimodal AI assessment, the comprehensive cultural value score of Pingjiang Historical District shows a clear distinction between the core and the periphery. The core protection area (accounting for 22% of the total area) carries 68.5% of the cultural value, among which the World Cultural Heritage sites such as the Couyuan Garden and the Quanjin Guild Hall form the peak value. Semantic analysis shows that the names of streets and lanes, river systems and traditional living customs are the core elements of community emotional identification. It is worth noting that image recognition has found that among traditional residences that are not cultural relics protection units, 17% have high potential value due to well-preserved wooden structures, brick carvings and other features. Such buildings are often overlooked in the current protection system but are actually crucial.

Based on the analysis of surface temperature and building energy consumption, it is found that the thermal stress in traditional dense blocks is significant in summer. The daily average maximum temperature in the dense area is 2.8°C higher than that in modern blocks, but the energy consumption per unit area is 18% lower, demonstrating the ecological value of traditional wisdom. Meanwhile, the vulnerability spatial analysis indicates that the ventilation corridor in the eastern region is interrupted due to river filling, and the heat island intensity is 1.5°C higher than that in the western region. The fire risk is prominent in the brick-and-wood mixed structure area in the north. Due to the aging drainage system in the central region, the risk points of urban flooding account for 43% of the entire area. Cross-analysis of environmental and semantic data further reveals that residents' perception of climate change (such as "more muggy summers") is highly correlated with the measured data ($r=0.76$, $p<0.01$), verifying the importance of native knowledge in vulnerability assessment.

4.4 Scenario Simulation and Comparison of Collaborative Optimization Strategies

Table 1. Comparison of Three Regeneration Scenarios Generated by Multi-objective Optimization

Scenario	Cultural Value Retention Rate	Carbon Reduction Rate	Key Intervention Measures	Applicable Subzones
Scenario A: Conservative Regeneration	96%	18%	Traditional craftsmanship restoration; passive cooling via alleyway ventilation optimization; rainwater permeable pavement	Core protected areas around couple's garden
Scenario B: Moderate Regeneration	90%	32%	Hidden rooftop PV; adaptive reuse of vacant buildings; AI-optimized waste management	General historic area with mixed functions
Scenario C: Comprehensive Regeneration	79%	48%	Green material replacement; sponge city facilities; energy-efficient retrofitting of building interiors	Marginal areas with lower cultural value

Based on the dual-objective optimization model, three types of update scenarios are generated for different partitions (Table 1). Scenario A (Conservative Update) is applicable to the core area of cultural value (such as the area around the decoupling garden). It follows the principle of "micro-intervention" and focuses on traditional restoration techniques and passive cooling (such as enhancing ventilation in the tunnels and restoring evaporative cooling of water bodies). The retention rate of cultural value reaches 96%, but carbon reduction is limited to 18%. Scenario B (Moderate Update) For general style areas, under the premise of maintaining the authenticity of the appearance, concealed energy-saving technologies (such as rooftop photovoltaic and rainwater recycling) are introduced. The cultural value is retained by 90%, and carbon reduction is increased to 32%. Scenario C (Comprehensive Update) : For marginal areas with lower value and prominent environmental issues, moderate renovations (such as internal function replacement and application of green building materials) are allowed. The cultural value is retained by 79%, but carbon reduction reaches 48%, and climate resilience is significantly enhanced (the risk of waterlogging is reduced by 52%).

The Pareto frontier analysis of multi-objective optimization indicates that there is no single optimal solution, but decision-makers can find a balance point within the range of 85% to 95% of cultural value retention. For instance, by adopting a strategy of focusing on Scenario B and partially integrating Scenario C in the area of style coordination, it is possible to retain 88% of cultural value while reducing carbon emissions by 37%. Sensitivity analysis further identified the key influencing factors. The material of building roofs (with an impact weight of 0.32 on the heat island effect) and the height-to-width ratio of streets and alleys (with an impact weight of 0.29 on ventilation efficiency) are the main leverage points for collaborative optimization.

5 Conclusions and Policy Recommendations

This research has achieved a breakthrough in the coordinated development of cultural heritage protection and sustainable renewal in historical urban areas by constructing a multimodal artificial intelligence system. The research results show that multimodal AI technology has successfully achieved methodological innovation in the transformation of cultural value from qualitative description to quantitative assessment. In the empirical research of the Pingjiang Historical District in Suzhou, this system not only accurately identified the officially protected cultural relics units, but also discovered the potential value buildings accounting for 17% of the total number of traditional residences. These buildings have high cultural value due to their well-preserved traditional architectural features, but they are often overlooked in the current protection system. This discovery provides an important scientific basis for improving the protection list of historical urban areas.

In terms of climate adaptability analysis, research has revealed the unique value of traditional architectural wisdom. Data analysis shows that although the summer thermal stress in traditional dense neighborhoods is 2.8℃ higher than that in modern neighborhoods, their energy consumption per unit area is actually 18% lower. This finding highlights the contemporary significance of the ecological wisdom of traditional architecture. Semantic analysis also found that residents have a high degree of recognition for traditional regulation methods such as ventilation in alleys and cooling of water bodies, indicating that the green renewal of historical urban areas should attach importance to the modern translation of traditional wisdom rather than simply applying new technical standards.

The dual-objective optimization model constructed in the research effectively resolves the trade-off dilemma between protection and update. The feasibility of the differentiated implementation path was proved by three update schemes generated through the multi-objective genetic algorithm. In particular, the moderate update plan achieves a 32% carbon reduction rate while maintaining 90% of cultural value, providing an ideal reference for most historical regions. Sensitivity analysis also identified key influencing factors such as the material of building roofs and the height-to-width ratio of streets and alleys, providing clear guidance for the priority order of update measures.

5.1 Policy Recommendations

Based on the research results, we suggest innovating the governance system of historical urban areas and establishing a planning and management mechanism oriented towards both culture and climate. Incorporate multimodal AI assessment into the legal procedures of historical urban area protection planning, and establish a dual evaluation system of cultural value assessment and climate adaptability assessment. Specifically, corresponding technical standards and planning guidelines can be formulated to clarify quantitative indicators and operational methods. At the same time, a dynamically updated digital archive platform for cultural heritage should be established to achieve regular and precise protection and monitoring.

At the implementation level, it is recommended to implement a precise update strategy of zoning and classification based on the AI assessment results. The historical urban area is divided into three policy zones: core protection area, style coordination area and renewal demonstration area, and differentiated renewal strategies are adopted respectively. The core protection area mainly focuses on the restoration of traditional techniques, with an emphasis on improving infrastructure. In the area of style coordination, adaptive green technologies are encouraged to be adopted while maintaining the authenticity of the appearance. The renewal demonstration zone can focus on exploring the integrated innovation of traditional wisdom and modern technology.

At the same time, it is necessary to establish a multi-party collaborative implementation mechanism involving the government, experts, communities and enterprises. Establish a collaborative platform to fully leverage the advantages of all parties. Among them, the government is responsible for policy guidance and supervision, research institutions provide technical support, community residents express their demands through participatory planning, and enterprises are responsible for technological innovation and project implementation. Special attention should be paid to community participation. AI semantic analysis technology systems can be utilized to collect and identify residents' demands. A historical urban area renewal fund can be established, and incentive measures can be adopted to encourage property owners to participate in protective renewal.

There are still some aspects in this study that need to be further improved. The cultural value assessment model still needs to be deepened, and more in-depth research is required on the spatial mapping relationship of intangible cultural heritage. Climate adaptability analysis also needs to incorporate longer-term meteorological forecast data to better cope with future climate change. The implementation effect of the multi-party collaboration mechanism also needs to be tested in practice.

Future research will focus on advancing in three directions. The first is to develop a predictive model for the evolution of cultural values over time to achieve dynamic protection. Second, research methods for assessing heritage resilience under extreme climate events to enhance disaster prevention capabilities. Third, explore a cultural asset management model that combines artificial intelligence and blockchain, and innovate protection and utilization mechanisms. Through these studies, the ultimate goal is to build a more intelligent, precise and inclusive methodology system for the sustainable development of historical urban areas, providing beneficial experience for the high-quality development of global urbanization.

Reference

1. Wang, K., & Fouseki, K. (2025). Sustaining the Fabric of Time: Urban Heritage, Time Rupture, and Sustainable Development. Land, 14(1), 193. https://doi.org/10.3390/land14010193. https://doi.org/10.3390/land14010193.
2. Li Aihua, Li Changen. Research on the Coupling of High-Quality Digital Economy Development and New Urbanization: An Empirical Analysis Based on 31 Provinces in China [J]. Times Economy, 2026, 23(8): 146-151. DOI: 10.19463/j.cnki.sdjm.2026.08.032.
3. Kong, L., Xu, X., Wang, W., Wu, J., & Zhang, M. (2021). Comprehensive evaluation and quantitative research on the living protection of traditional villages from the perspective of "Production–Living–Ecology". Land, 10(6), 570. https://doi.org/10.3390/land10060570.
4. Blagojević, M. R., & Tufegdžić, A. (2016). The new technology era requirements and sustainable approach to industrial heritage renewal. Energy and Buildings, 115, 148-153. 10.1016/j.enbuild.2015.07.062.
5. Wamsler, C., Brink, E., & Rivera, C. (2013). Planning for climate change in urban areas: from theory to practice. Journal of Cleaner Production, 50, 68-81. DOI：10.3390/su6031359.
6. Fan Wei, Zeng Lei. Analysis of the Intelligent Data Generation Path for the Revitalization and Utilization of Cultural Heritage in the AI Era [J].Fan Wei, Zeng Lei. Analysis of the Intelligent Data Generation Path for the Revitalization and Utilization of Cultural Heritage in the AI Era [J]. Chinese Journal of Library Science, 2024, 50(2): 4-29. DOI: 10.13530/j.cnki.jlis.2024010.
7. Chen Peiyang, Zhou Jiangyi, Wu Peiya. Research on the Update Strategies of Industrial Remnants in Old Urban Areas of Large Cities from the Perspective of Innovative Urban Areas [J]. China Famous Cities, 2024, 38(3): 25-30, DOI: 10.19924/j.cnki.1674-4144.2024.003.004.
8. Zhou Jianyun, Jiang Dingzhe, Bao Ziting, et al. Evaluation of Historical Landscape Characteristics - Linking Protected Areas with Urban Historical Landscapes [J]. International Urban Planning, 2023, 38(2): 124-132. DOI: 10.19830/j.upi.2020.149.