coatings

MDPI

Article

Surface Damage Detection and Analysis for Reduction-Fired Cyan Square Bricks in Jiangnan Gardens via YOLOv12

Lina Yan [1], Yile Chen [2], Xingkang Jia [1] and Liang Zheng [2,*]

[1] College of Arts, Shanghai Zhongqiao Vocational and Technical University, No. 3888 Caolang Road, Jinshan District, Shanghai 201514, China; linayan@shzq.edu.cn (L.Y.); jiaxingkang@shzq.edu.cn (X.J.)
[2] Faculty of Humanities and Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau 999078, China; chenyile@must.edu.mo
* Correspondence: zhengliang@must.edu.mo

Abstract

As an outstanding UNESCO World Heritage Site, the Jiangnan gardens feature both exquisite and fragile components. Reduction-fired cyan square bricks, serving as crucial paving materials, are long-term exposed to natural and anthropogenic factors, making them prone to various types of surface damage and urgently requiring efficient, non-destructive detection methods to support scientific conservation. Traditional manual inspection methods suffer from low efficiency, strong subjectivity, and potential disturbance to the fragile heritage structures. This study focuses on developing an intelligent detection method based on advanced computer vision, employing the YOLOv12 object detection model to achieve non-contact, automated identification of typical tile surface damage types in the Jiangnan gardens (such as cracking, stains, water stains, and wear). A total of 691 images of reduction-fired cyan square bricks collected on-site were used as training samples. The main conclusions of this study are as follows: (1) By constructing a dataset containing multiple samples and multiple scenes of reduction-fired cyan square brick images in Jiangnan gardens, the YOLOv12 model was trained and optimized, enabling it to accurately identify subtle damage features under complex texture backgrounds. (2) Overall indicators: Through the comparison of the confusion matrices of the four key training nodes, model C (the 159th epoch, highest mAP50–95) has the most balanced overall performance in multiple categories, with an accuracy of 0.73 for cracking, 0.77 for wear, 0.60 for water stain, and 0.65 for stains, which can meet basic detection requirements. (3) Difficulty of discrimination: Compared with stains and water stains, cracking and wear are easier to distinguish. Experimental results indicate that the detection method is feasible and effective in identifying the surface damage types of reduction-fired cyan square bricks in Jiangnan gardens. This research provides a practical and efficient "surface technology" solution for the preventive protection of cultural heritage, contributing to the sustainable preservation and management of world heritage.

Keywords: world heritage site; YOLOv12; reduction-fired cyan square bricks; surface damage

check for updates

Academic Editor: Alexander Tolstoguzov

Received: 19 August 2025
Revised: 7 September 2025
Accepted: 10 September 2025
Published: 11 September 2025

Citation: Yan, L.; Chen, Y.; Jia, X.; Zheng, L. Surface Damage Detection and Analysis for Reduction-Fired Cyan Square Bricks in Jiangnan Gardens via YOLOv12. *Coatings* **2025**, *15*, 1066. https://doi.org/10.3390/coatings15091066

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The Suzhou Historic and Cultural City Conservation Area, located in Suzhou City, Jiangsu Province, in China's Yangtze River Delta, boasts a history spanning over 4000 years, dating back to the earliest recorded history [1,2]. The ancient city of Suzhou, covering approximately 14.2 square kilometers, has remained unchanged for over 2500 years [3,4].

The Suzhou Historic and Cultural District boasts numerous historical sites, among which the Jiangnan gardens are the most representative [5]. In 1997 and 2000, six exemplary classical gardens within the Suzhou Old Town area, including the Humble Administrator's Garden, the Lingering Garden, and the Master of the Nets Garden, were inscribed on the UNESCO World Heritage List [6], becoming a treasure shared by all mankind.

Reduction-fired cyan square bricks, an integral part of the Jiangnan classical gardens, a World Cultural Heritage site, have a history of use dating back to the Song Dynasty (960–1279) [7–9]. Large-scale square bricks were primarily used in palaces, temples, and the halls of wealthy households. During the Ming and Qing Dynasties (1368–1840), the flourishing development of Jiangnan gardens led to an increase in use of reduction-fired cyan square bricks in flooring, making them an essential element of garden floor decoration. From the mid-to-late Qing Dynasty to the Republic of China (1840–1949), while traditional craftsmanship continued, hand-fired cyan square bricks gradually became rare [10,11]. Existing square cyan brick paving in Jiangnan classical gardens largely retains the original bricks [12]. These bricks, passed down through centuries, have witnessed the rise and fall of Jiangnan gardens and their evolution and are now a unique and indelible part of the World Heritage list.

Scholars have made a series of explorations on the brick-related research of Suzhou's architectural heritage. Li et al. used the constant humidity and constant temperature method of a saturated saline solution to measure the isothermal moisture absorption curves of Wujiang bricks in the 1980s and the late Qing Dynasty and gave a fitting formula in the humidity range of 5%−92% [13]. Wang et al. conducted physical and chemical analyses of the clinker bricks used in Soochow University using X-ray diffraction (XRD), Raman spectroscopy, scanning electron microscopy–energy-dispersive spectroscopy (SEM-EDS), and a full immersion test and studied the physical and chemical properties of bricks at different positions on the facade of the Smart Memorial Gymnasium [14].

There are also studies related to the material properties of bricks. Chun et al. studied ancient white bricks in the Yichun region from the perspective of material properties [15]. Li et al. investigated the cross-validation of the hygrothermal performance and isothermal adsorption experiments of historical Chinese gray bricks [16]. Zhao et al. considered Jiangsu, Zhejiang, and Anhui Provinces as sample areas and traced the ancient production processes for traditional Chinese gray bricks. They compared the porosity and water absorption of traditional gray bricks and modern antique-style bricks [17]. Zhang et al. focused on improving the performance of historical gray bricks made using an amide polymer [18].

However, Suzhou has a vast heritage site and uses a wide variety of classical hand-made materials, which existing research has not been able to fully cover, especially in the field of reduction-fired cyan square bricks. Previous studies mainly focused on the material elements, composition, and material properties of bricks. These studies required on-site sampling of the bricks to conduct element and composition detection. To obtain sufficient sample sizes, this sampling research method usually involved scraping or fragment sampling from the original bricks. This inevitably damaged the integrity of the original bricks.

With the emergence of deep learning technology, more scholars hope to use these cutting-edge technologies to achieve non-contact inspection and identification of heritage surfaces. In 2024, our research team conducted surface damage recognition and object detection using YOLOv4 on shedthin tiles on the roofs of Suzhou's traditional garden architecture [19]. Chen et al. combined drone technology with deep learning to perform pathological detection, focusing on five categories, namely material loss (ML), discoloration and deposition (DD), cracks (CRs), surface spalling (SS), and biological invasion (BI), at

Soochow Hospital [20]. Despite this, we still need to pay more attention to this precious world heritage and expand more research combining materials science and heritage studies. Among the numerous deep learning models, Ultralytics Inc. (Frederick, MA, USA) released YOLOv12 in 2025 [21–23]. This model introduced an attention-centric architecture that departed from the traditional CNN approach used in previous YOLO models while maintaining the real-time inference speed essential for many applications. This model achieved state-of-the-art object detection accuracy while maintaining real-time performance through innovative approaches to the attention mechanism and overall network architecture. This also ushered in new applications for deep learning technology in the field of surface damage analysis for architectural heritage buildings.

In contrast to prior research on brick heritage, which has predominantly focused on chemical composition and microscopic properties, studies specifically addressing the surface damage of reduction-fired cyan square bricks remain scarce. The integration of intelligent detection technologies in this domain is even more limited. This study pioneers the application of an advanced object detection model, YOLOv12, for the non-contact and highly efficient identification of surface damage on these bricks. The primary purpose is to develop an automated approach that transcends the limitations of traditional manual inspection, which is often subjective and time-consuming. This research not only demonstrates the feasibility of using deep learning for heritage material diagnostics but also provides a robust and scalable framework for the condition assessment of garden heritage, thereby offering a significant departure from and contribution to the existing body of work.

2. Materials and Methods

2.1. Research Object: Reduction-Fired Cyan Square Bricks in Jiangnan Gardens

Reduction-fired cyan square bricks, a crucial paving material in Jiangnan gardens, play a crucial role in the garden's overall layout. These bricks, despite centuries of use, largely retain their original paving material and layout, having never undergone large-scale replacement. They bear witness to the garden's historical beauty [24,25]. However, compared to other components of ancient buildings, such as beams, columns, and window lattices, these bricks have received significantly less protection [26,27]. Today, due to multiple factors, including climatic changes (such as prolonged rainwater erosion and drastic temperature fluctuations), frequent human traffic, and accidental impacts, many bricks have developed varying degrees of cracking, wear, hollowing, and even loss.

Scientific and effective surface inspection measures are a prerequisite for protecting these reduction-fired cyan square bricks. Precise inspection techniques can comprehensively capture subtle surface damage (such as early cracks and localized wear), thereby promptly identify potential problems and buy time for subsequent conservation efforts. If left undetected and allowed to develop, minor damage can escalate into major destruction, ultimately leading to the irreversible loss of cultural heritage. With such consequences in mind, this study specifically focuses on the floor tiles of six Jiangnan classical gardens, a World Heritage Site, within the ancient city of Suzhou. It systematically examines surface damage on reduction-fired cyan square bricks. Its core goal is to utilize scientifically effective inspection methods to promptly identify various damage issues, providing a solid foundation for developing targeted conservation strategies to safeguard these precious World Heritage Sites (Figure 1).

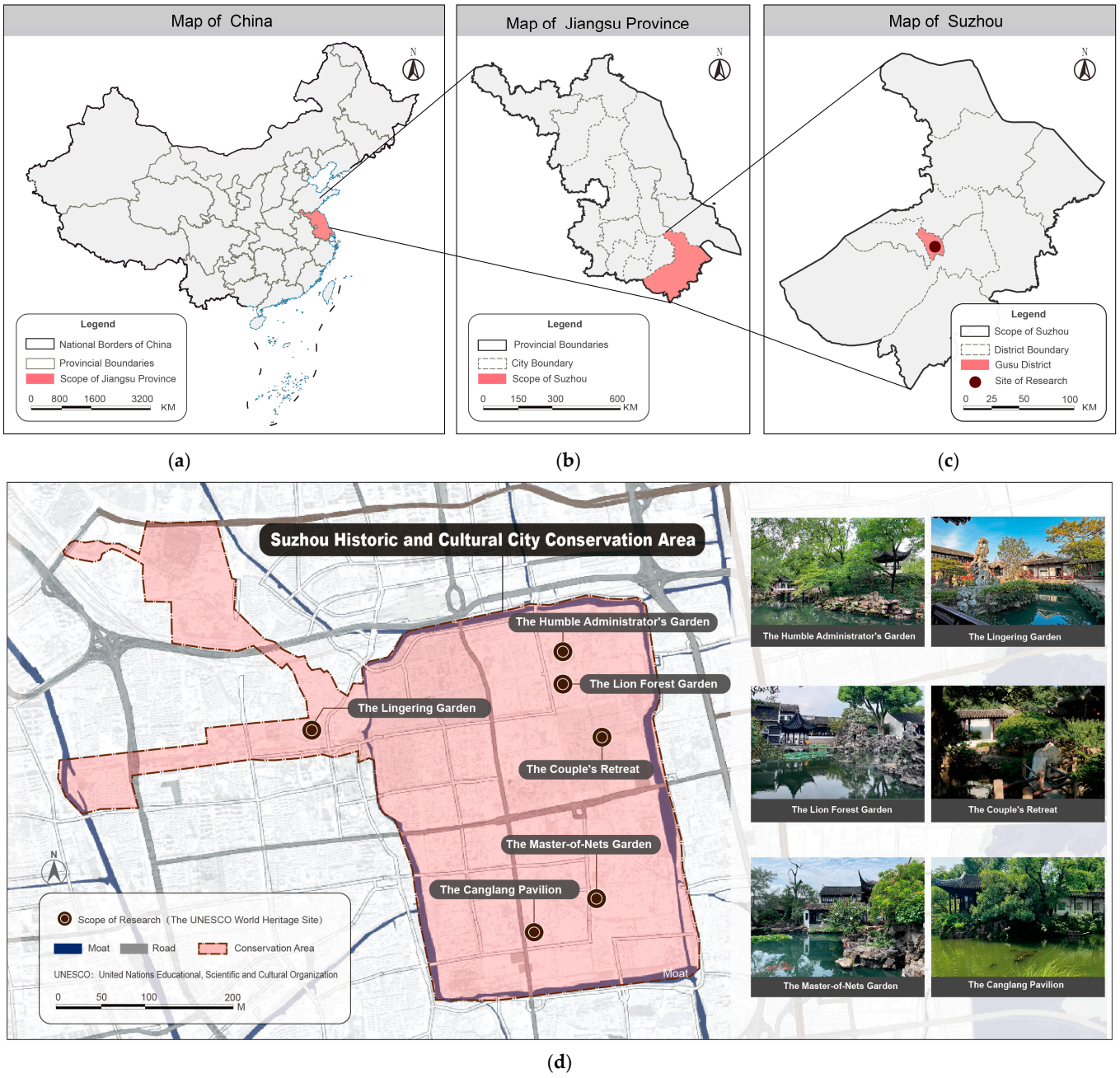


Figure 1. Research area: Suzhou Historic and Cultural City Conservation Area in Suzhou. (**a**) Location of Jiangsu Province; (**b**) location of Suzhou; (**c**) site of research; (**d**) research scope of the Suzhou Historic and Cultural City Conservation Area.

Reduction-fired cyan square bricks are used specifically for paving the ground in Jiangnan classical gardens. Their surface is flat, smooth, and delicate, and they are widely used in the ground base of the halls and corridors of the main buildings in the gardens (Figure 2).

The specifications of reduction-fired cyan square bricks vary depending on the specific landscape and design requirements. Outdoor paving uses reduction-fired cyan square bricks primarily for relatively regular courtyards, terraces, and paths. Typical side dimensions range from 20 cm to 40 cm, with thicknesses ranging from 6 cm to 10 cm. Indoor living rooms use reduction-fired cyan square bricks, which are more delicate and strive for a smooth and glossy surface. They are typically larger and thinner than outdoor bricks,

with typical side dimensions ranging from 40 cm to 70 cm and thicknesses ranging from approximately 6 cm to 10 cm.

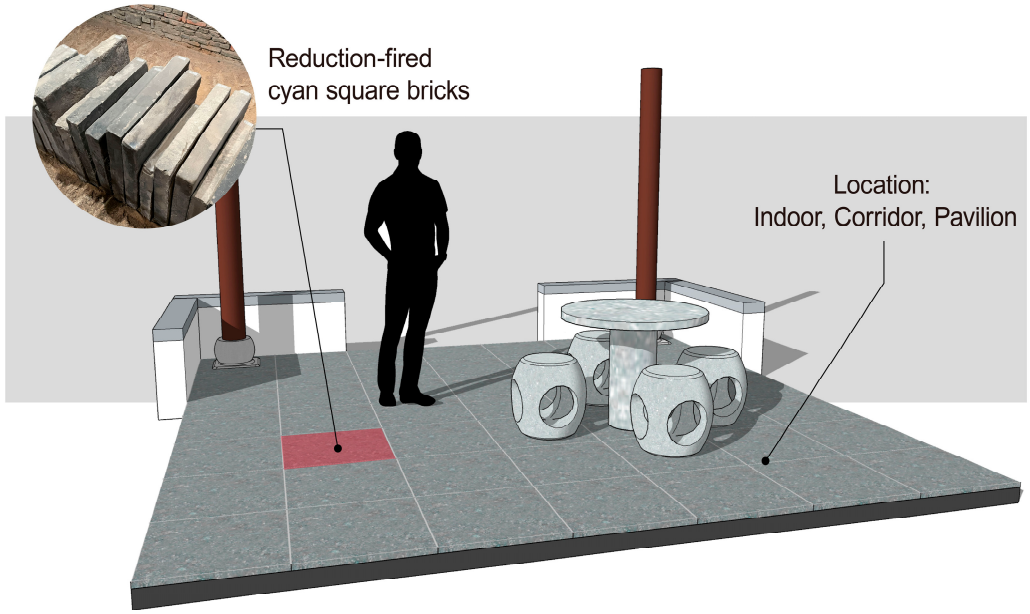


Figure 2. The laying location of reduction-fired cyan square bricks.

The laying of reduction-fired cyan square bricks requires extremely stringent craftsmanship, with each brick embodying the wisdom and ingenuity of ancient gardeners. Traditional brickmaking involves carefully selecting high-quality clay, repeatedly treading to remove air bubbles, and then undergoing a lengthy shade-drying process, high-temperature firing, and a water-reduction process to achieve the cyan color. Finally, they undergo rigorous inspection to eliminate inferior products. The entire brick production process is lengthy and characterized by a high loss rate, making it difficult to replicate with other flooring materials. Today, these traditional handmade reduction-fired cyan square bricks can be found in several historic buildings in the Jiangnan region, including the main pavilion in the Humble Administrator's Garden and the corridors of the Lingering Garden (Figure 3).



(**a**) (**b**) (**c**)

Figure 3. Reduction-fired cyan square bricks are commonly used in traditional gardens. (**a**) The Hall of Ten Thousand Volumes in Suzhou's Master of the Nets Garden; (**b**) inside the Humble Administrator's Garden Pavilion; (**c**) Inside the long corridor in the Lingering Garden.

2.2. Analysis of Damage Types and Factors

Ancient reduction-fired cyan square bricks are a distinctive material used in the floor paving of traditional Jiangnan buildings. However, these materials deteriorate over

time, leading to a decrease in their physical and mechanical properties and durability, compromising the usability of the floor. These historic buildings, which have survived for hundreds or even nearly a thousand years, have undergone numerous restorations, but the traditional firing techniques are rapidly disappearing, and high-quality, old-style reduction-fired cyan square bricks are now extremely rare. If these bricks, which carry this historical context, were to disappear completely, the floorscape of Jiangnan's traditional architecture would lose its soul, and the integrity and authenticity of this World Heritage Site would suffer irreparable damage. Therefore, their protection and restoration are time-sensitive (Figure 4).

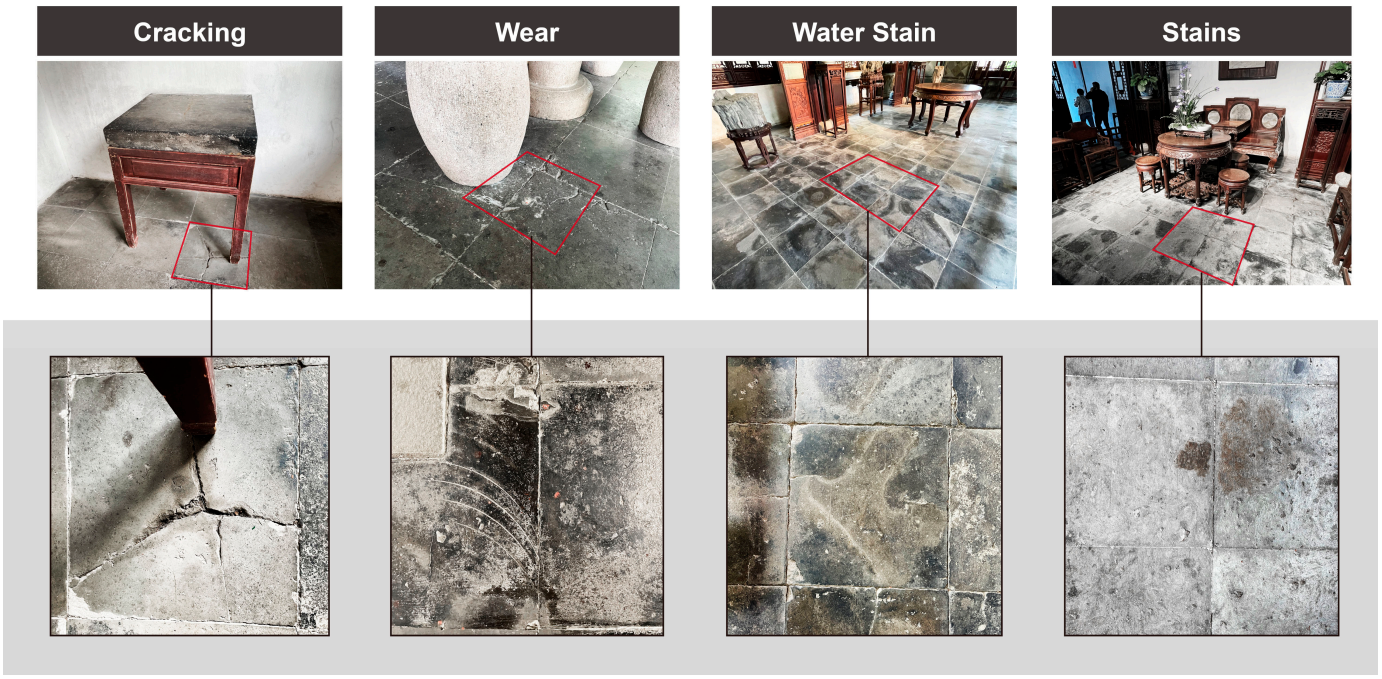


Figure 4. On-site photos of the damage to reduction-fired cyan square bricks.

Based on on-site inspections of multiple gardens, the research team found that the surface damage of reduction-fired cyan square bricks in Suzhou's world heritage ancient buildings was quite significant, and restoration work was time-sensitive. The damage to reduction-fired cyan square bricks can be mainly attributed to two factors, climate factors and human factors, which also make protection work more challenging [28,29]. This includes the following: (1) In terms of climate, its impact is closely related to the material properties of the tiles themselves. The Jiangnan region is humid and rainy, and reduction-fired cyan square bricks are exposed to high humidity for a long time. Traditional reduction-fired cyan square bricks have high porosity, and rainwater can easily penetrate deep into the brick body through capillary action. Long-term retention of moisture will cause irregular dark patches on the brick surface; at the same time, the humid environment will promote the growth of mold and accelerate the peeling of the brick surface. (2) In terms of human use, the damage to reduction-fired cyan square bricks is mostly caused by improper behavior during daily use. Surface contamination such as tea stains, oil stains, and paint, as well as cracking caused by heavy pressure, can directly damage reduction-fired cyan square bricks. We can identify four common types of damage based on the two main causes mentioned above: wear, cracking, water stains, and stains (Figure 5).

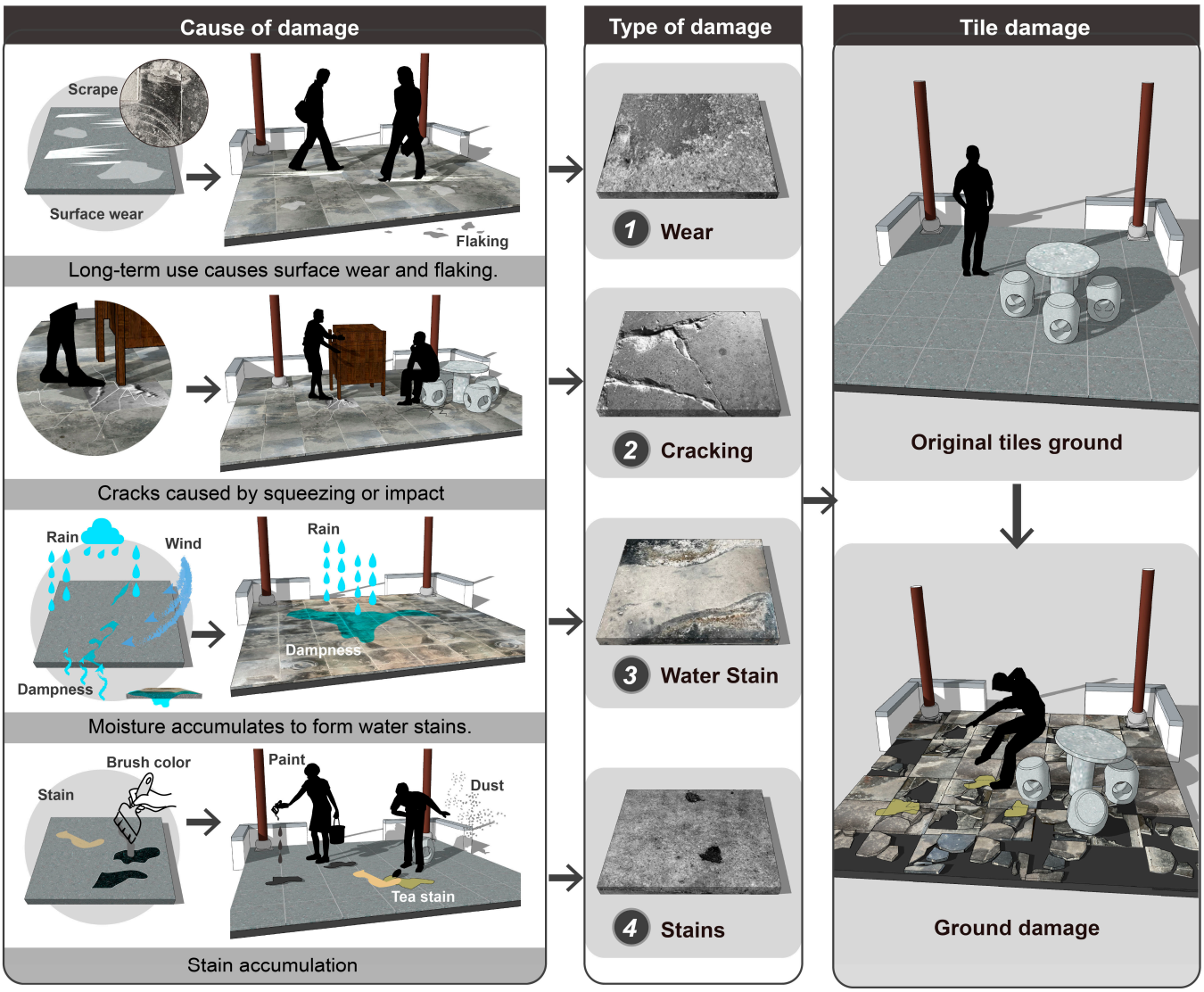


Figure 5. Causes and processes of damage of reduction-fired cyan square bricks.

2.2.1. Wear

Many factors affect the damage of reduction-fired cyan square bricks. From the perspective of the material itself, there are certain process defects: The firing temperature of traditional reduction-fired cyan square bricks is relatively low (about 1000 °C). The porosity is high [30–32]; it easily absorbs moisture and pollutants, and the surface is more susceptible to damage from the external environment. In terms of weather factors, the Jiangnan region has abundant rain, and long-term rain and moisture have a significant impact. The surface mortar will gradually wear out and even cause the surface to peel off. In addition, poor air flow inside the building, combined with moisture and frequent walking, makes the surface peeling or wear of these reduction-fired cyan square bricks on the indoor floor more serious. The rise in tourism has sped up surface wear.

2.2.2. Cracking

Reduction-fired cyan square bricks may crack or even break due to compression or impact. The Lake Tai basin, where the gardens are located, primarily consists of soft soil. This type of soil structure is less stable, and under the influence of continuous rainwater infiltration, the foundation is prone to uneven subsidence. When the foundation settles slightly, the reduction-fired cyan square bricks laid on top will be pulled due to unbalanced

forces, resulting in cracks. In addition, the temperature in the Jiangnan region often exceeds 35 °C during the hot summer season, and the reduction-fired cyan square bricks will expand due to heat. Because the reduction-fired cyan square bricks are laid closely and constrained by the surrounding structure, their expansion space is limited, resulting in the edges of the bricks squeezing each other and causing edge cracking. The reduction-fired cyan square bricks in Suzhou's gardens have experienced a century of daily use. Certain heavy objects or the movement of furniture will inevitably cause impact and compression damage to the floor tiles. These external impacts and compressions will directly act on the surface and internal structure of the reduction-fired cyan square bricks, causing serious cracks in the tiles. Excessively large cracks will not only accelerate the fragmentation of the bricks themselves but will also affect walking and even cause tripping due to cracking and warping.

2.2.3. Water Stain

The average annual humidity in Jiangnan is >75%, and the rainy season is continuous (such as the plum rain season) [33,34]. The porosity of traditional reduction-fired cyan square bricks is relatively high, and rainwater can easily penetrate deep into the brick body through capillary action. When water remains in the brick body for a long time, the surface of the reduction-fired cyan square bricks will not have regular dark patches, resulting in blurred boundaries. In addition, after some water stains evaporate, white salt crystals will precipitate at the edge of the water stains on the brick surface, forming a white ring of alkali. After repeated dry–wet cycles, the water stains carry a soluble salt into the pores of the brick body. When drying, the crystals precipitated by the salt will expand, squeezing the pores inside the brick body and causing the brick structure to crack and peel off. It can be said that water stains are the initial manifestation of cracks and damage in reduction-fired cyan square bricks. The three factors are closely related.

2.2.4. Stains

Stains, unlike water stains, are primarily caused by human use. Stains primarily include tea stains, food grease, and paint (such as restoration paint and lime mortar). While these stains do not directly damage the surface of the bricks, if left untreated over time, they will gradually penetrate the rough, porous surface or cracks of the bricks. Grease stains can cause the surface of the tiles to become sticky and rigid, accelerating dust adhesion. High humidity enables pollutants to form sols, which increase their permeability and lead to the development of carbon-black pits on the surface of reduction-fired cyan square bricks. This damage to the bricks is irreversible.

Sample data collected through on-site surveys indicates that surface damage to the reduction-fired cyan square bricks within Suzhou's world heritage buildings is particularly prominent, necessitating urgent restoration. For historical buildings, restoration efforts must preserve the original features to maximize authenticity and completeness while also ensuring the continued function of the building. Such an outcome significantly increases the difficulty of protecting and repairing the reduction-fired cyan square bricks. If timely protective measures are not implemented, damage to the reduction-fired cyan square bricks will worsen, affecting not only the building's overall appearance but also its functionality. To promote intelligent and sustainable conservation of world heritage buildings, this study utilizes machine learning technology and a contactless identification method to establish a real-time, rapid monitoring system for surface damage to reduction-fired cyan square bricks.

2.3. Image Collection Source

The image samples for this study were selected from the floors in buildings and pavilions within Suzhou's six major classical gardens: the Humble Administrator's Garden, the Lingering Garden, the Master of the Nets Garden, the Canglang Pavilion, the Lion Grove Garden, and the Couple's Retreat Garden. To ensure the authenticity and representativeness of the samples, researchers conducted a systematic field survey between May and July 2025. Given the critical impact of lighting on image quality, sample collection was conducted on sunny days with good lighting conditions. High-definition cameras were used throughout the process to maximize the preservation of detailed features of surface damage on the reduction-fired cyan square bricks.

During the three-month survey, the research team collected a total of 977 samples of reduction-fired cyan square bricks through repeated on-site visits and meticulous regional inspections. To ensure the reliability of the data for subsequent research, these samples underwent a rigorous screening process. After individually screening and eliminating invalid samples with problems such as overexposure, image blur, and excessive angle deviation, the researchers ultimately identified 691 valid samples. Among the valid samples from various gardens, the Humble Administrator's Garden accounted for the largest proportion, with 309 samples. This result is due to the garden's large scale, the widespread distribution of its tiles, and the severe damage they suffered. The Lingering Garden and the Lion Grove Garden each had 70 valid samples. The Master of the Nets Garden had 84 samples. The Canglang Pavilion had 81 samples, and the Couple's Retreat Garden had 77 samples (Figure 6).

UNESCO ID	Name & Location		Coordinates	Property	Acquisition time	Quantity
813–001	The Humble Administrator's Garden		N31 19 27.721 E120 37 44.62	5.195 ha	05/15/2025 07/24/2025 07/26/2025	309
813–002	The Lingering Garden		N31 18 56.2 E120 35 33.515	2.331 ha	07/24/2025 07/26/2025	70
813–003	The Master-of-Nets Garden		N31 17 52.66 E120 38 2.418	0.54 ha	05/15/2025 07/24/2025	84
813-005	The Canglang Pavilion		N31 17 41.348 E120 37 30.972	1.174 ha	07/24/2025 07/26/2025	81
813-006	The Lion Grove Garden		N31 19 15.91 E120 37 44.191	0.874 ha	05/15/2025 07/24/2025	70
813-008	The Couple's Retreat Garden		N31 18 58.324 E120 38 19.028	0.789 ha	05/15/2025 07/24/2025	77

Figure 6. Image acquisition source and quantity.

2.4. Methodology

2.4.1. Research Process

This study focused on the surface damage of reduction-fired cyan square bricks in Jiangnan classical gardens. Using YOLOv12 object detection technology, we automatically identified and classified damaged images and proposed an efficient and accurate solution for analyzing ground damage. This research process consisted of six steps, as shown in Figure 7.

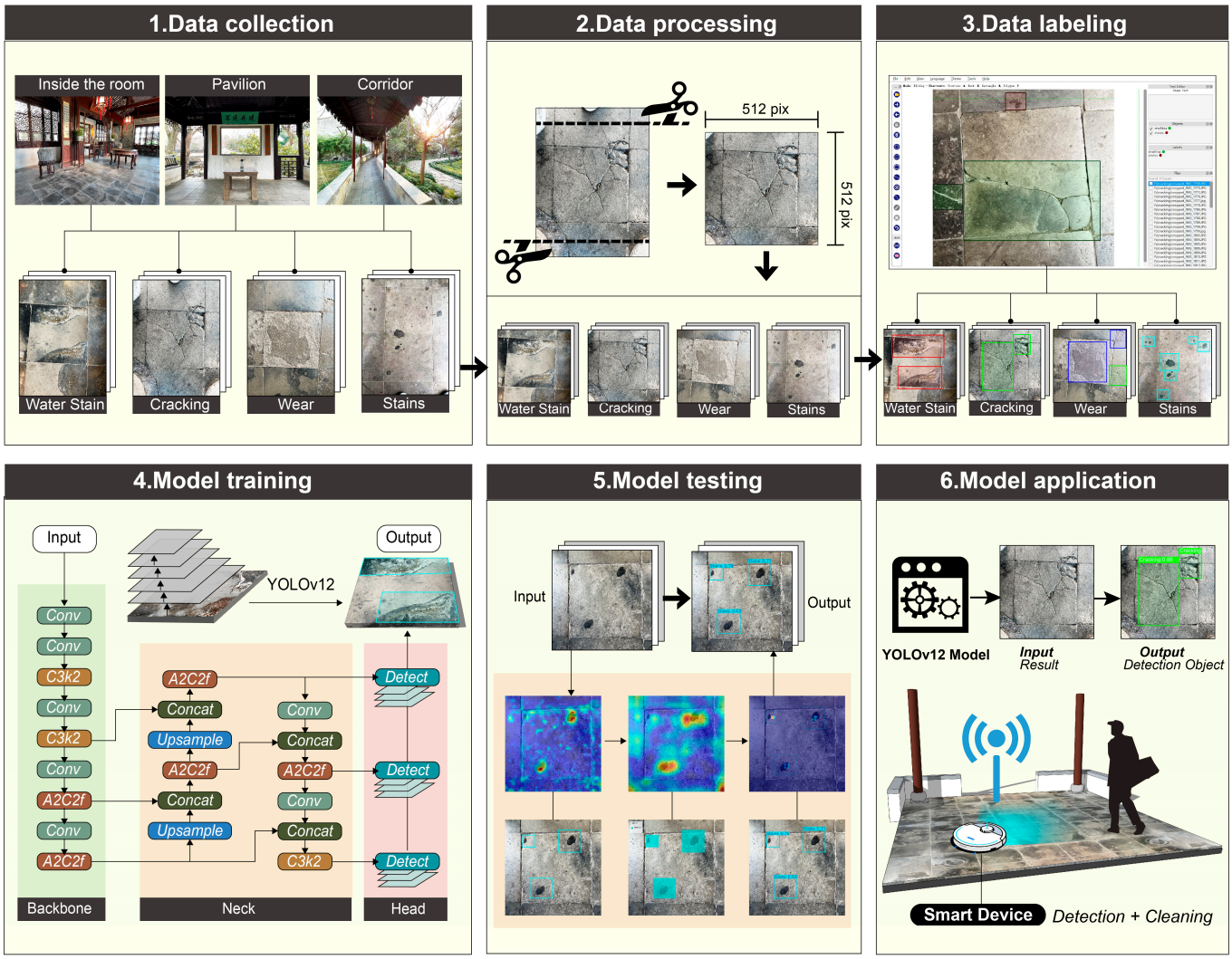


Figure 7. Research methods and processes.

(1) Data collection: To ensure the effectiveness of the damage-type identification model for the surface of reduction-fired cyan square bricks, this study strictly standardized the sample collection process. The researchers used six classical gardens in Suzhou as sampling points and entered typical scenes such as indoors, pavilions, and corridors to take photos. These samples included major damage types, such as water stains, cracks, wear, and stains. During the collection process, the researchers paid attention to shooting variables such as lighting, angle, and distance to ensure the diversity and representativeness of the data, thereby obtaining complete damage image data. After three months of multiple field shootings, a total of 977 damage samples were obtained.

(2) Data processing: To ensure the effectiveness of YOLOv12 model training, the images need to be standardized. The researchers adjusted the size of all collected images to 512 × 512 pixels and set the resolution to 96 dpi. These settings can eliminate the

interference of image size, clarity, and other issues on model training from the source. This allows each sample to participate in training under the same technical parameters. In addition, the researchers screened each unqualified sample for blurry images caused by instantaneous lens shaking during the shooting process and color deviation caused by light changes. The researchers deleted invalid images, and for valid samples missing in key scenes, professionals were arranged to reshoot to ensure the integrity of the sample set. After systematic shooting, layer-by-layer screening, and data processing, the researchers finally identified 691 valid reduction-fired cyan square brick photos as machine learning training samples.

(3) Data labeling: To enable the model to accurately identify the type and specific location of damage to reduction-fired cyan square bricks, the research team used the AnyLabeling professional labeling tool to perform detailed manual data labeling on the floor tiles. The labeling work was completed by several architects and cultural relics experts. The researchers carefully observed each pre-processed image; accurately located the damage for four typical types of cracking, stains, water stain, and wear; and generated standard labels. Each sample has a corresponding damage type and a specific location. After the initial labeling, the team conducted three sample labeling checks and proofreading to ensure the accuracy of the labeling.

(4) Model training: This study trains the surface damage detection of reduction-fired cyan square bricks based on the YOLOv12 framework. Following the YOLOv12 model training framework, the labeled data is input into it for training. YOLOv12 uses the backbone (backbone network), neck (neck network), and head (head network) training links [35,36]. The backbone (backbone network) consists of the Conv (convolution layer), C3k2, and A2C2f modules. C3k2 is used for auxiliary feature extraction, and A2C2f is used for feature extraction, further processing, and refining features. The neck (neck network) Contains the Concat (splicing layer), Upsample (upsampling layer), and A2C2f modules. The features extracted by the backbone are fused and adjusted. Through upsampling and splicing operations, feature information at different levels is integrated to enhance feature expression. The head network consists of multiple detection layers, which are responsible for the final target detection task and output the category and location information of the detected target. Please refer to Appendix A for the operating environment parameters of the model.

(5) Model testing: To verify the effect of model training, the researchers tested the trained model. The overall verification indicators include precision, the recall rate, mAP@0.5, mAP@0.5:0.95, and the comprehensive fitness index (Fitness). To evaluate the model's detection effect on specific damage types, researchers drew the model's precision–recall curve and calculated the mAP@0.5 for each category. It can intuitively reflect the changes in precision and recall rate under different confidence thresholds, revealing the detection difficulty and model adaptability.

(6) Model application: This study proposes a ground intelligent detection application model that takes into account the unique properties of the reduction-fired cyan square bricks in Jiangnan classical gardens and the actual needs of protection work. Its technical path is to embed the trained algorithm model into a ground intelligent detection terminal with a cleaning function, making the terminal an intelligent detection system that integrates technology for identifying ground damage and cleaning functions. In this way, an intelligent detection device that utilizes technology for identifying ground damage and has a cleaning function is created. In actual operation, it can enable garden management personnel to protect floor tiles more conveniently. Through the intelligent system, the status of reduction-fired cyan square bricks can be monitored in real time, and targeted repair and maintenance plans can be formulated.

2.4.2. YOLOv12 Model Structure Design

This research is based on the YOLOv12 model structure, which enables robust, real-time, and high-precision identification of subtle surface damage to floor square bricks in complex garden environments. Jiangnan gardens are characterized by high humidity, complex textures, and significant lighting variations. Compared to traditional manual inspections, the convolution–attention hybrid nondestructive inspection model can automatically screen for various surface defects, including cracks, stains, water stains, and wear, without disturbing the heritage site itself. It is particularly effective in detecting low-contrast, small-scale objects and interfering with background textures.

YOLOv12's overall structure follows the classic three-part backbone, neck, and head but introduces enhanced feature modeling and cross-scale fusion units at key locations. The backbone extracts multi-scale texture and structural features, while the neck aggregates semantics and details at different resolutions through top-down and bottom-up approaches. The head performs joint classification and regression predictions on multi-scale feature maps. As shown in Figure 8, the model starts with an input ($512 \times 512 \times 3$) and undergoes multiple strided convolutions and residual block downsampling to form a pyramidal feature representation from high-resolution details to low-resolution semantics. Upsampling, concatenation, and attention enhancement are then used to enable information flow between scales, ultimately outputting multi-scale detection results for four types of damage identification. The following are the training settings for the three key components in the model structure. Please refer to Appendix B for the structural parameters of the model.

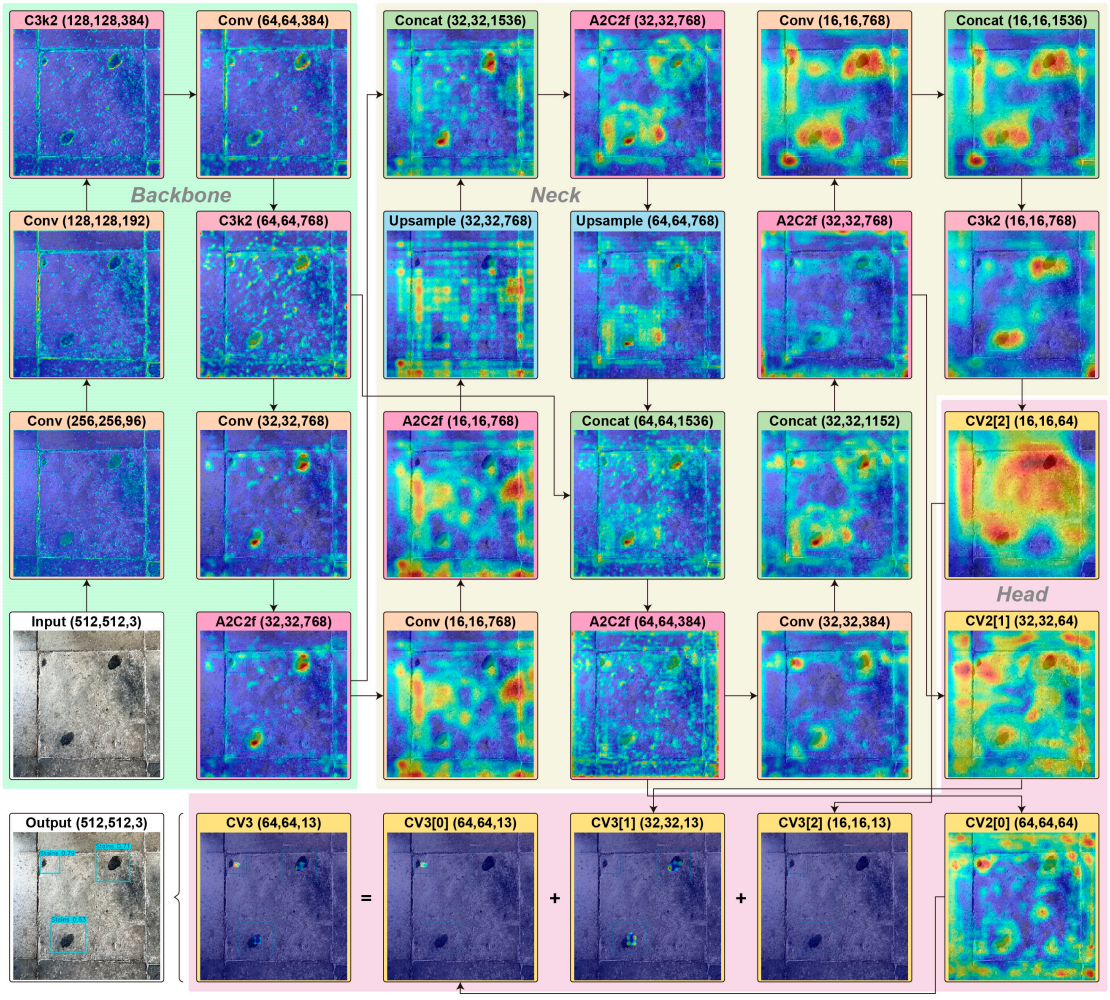


Figure 8. YOLOv12 model structure design.

(1) The backbone structure employs Conv and C3k2 stacking, which enhances the preservation of details and the transfer of gradients on complex brick surface textures. In the studied model, the input is subjected to two stride convolutions to obtain scale representations such as (256, 256, and 96) and (128, 128, and 192). Subsequently, C3k2 residual-style aggregation (such as (128, 128, and 384), (64, 64, and 768), etc.) is used to suppress invalid textures through group convolution and layer-by-layer fusion while retaining key clues such as crack edges and wear pits. In the deeper layer, A2C2f and its internal AAttn are introduced, and local attention and lightweight MLP are superimposed on the convolution backbone to highlight the discriminant signal of suspected diseased areas in low-contrast and noisy backgrounds.

(2) The neck combines Upsample, Concat, A2C2f, and C3k2 to construct a multi-scale representation that is more friendly to small targets and fine textures. As the model reflows from the high semantic layer, it injects edge and texture details from the shallow layer into the fusion pipeline. The sequence from Upsample and Concat to A2C2f significantly improves the distinguishability of fine cracks, multi-morphological water stain edges, and scattered oil stains.

(3) The head uses multi-branch convolution (CV2/CV3) for parallel prediction at multiple scales and combines DFL (distribution focal loss) to improve the fine-grained accuracy of bounding box regression. In the model of this study, the detection heads of the three scales are all composed of a series of deep and shallow groupable/depthwise separable convolutions followed by a 1×1 output layer. Classification and regression are collaboratively optimized on the same feature map; DFL improves boundary regression from point estimation to distribution estimation, which is more friendly to targets with irregular shapes and lengths/thicknesses across the grid, such as fine cracks, missing corners, and worn edges.

In the real-world environment of Jiangnan gardens, objects often exhibit low-contrast, small-scale, and non-rigid shapes, while previous models (e.g., YOLOv5 and YOLOv8) struggle with the rich textures and crack patterns in the background. This model utilizes attention enhancement with A2C2f + AAttn, efficient modeling through grouping and depthwise separable convolutions, and a neck design that employs multiple upsampling concatenations followed by attention and residual reinforcement. The above technique allows the model to achieve enhanced discrimination of stains, shadows, and material noise while minimizing misjudgments of cracks, seams, and pavement lines, all while maintaining real-time performance. In contrast, the backbone and neck of YOLOv5 are more traditional, limiting their expressiveness on subtle boundaries and heterogeneous backgrounds. While YOLOv8 has evolved in its architecture, its default fusion and attention configurations remain insufficient for scenarios like this, where low-contrast, fine objects are combined with strong texture interference. Therefore, without sacrificing inference speed, YOLOv12 provides a more robust detection basis for multiple types of surface damage on reduction-fired cyan square bricks.

3. Results

3.1. Model Training Process

This study trained a model for detecting surface damage in reduction-fired cyan square bricks within the YOLOv12 framework. The model training parameters for this study are shown in Table 1 below. Training was performed based on the pretrained YOLOv12x weights. CUDA was enabled, using a 512×512 input size and a batch size of 16 for 500 epochs. The SGD optimizer with a cosine annealing learning rate strategy was used. Momentum was set to 0.937, and weight_decay was set to 5×10^{-4} to stabilize convergence and suppress overfitting. Warmup_epochs = 5 was used to mitigate initial gradient oscilla-

tion. AMP mixed precision was enabled with 16 workers to improve data pipeline and computational efficiency. Validation was enabled throughout the training process to monitor mAP50–95, precision, and recall. In terms of parameter exploration, we tried AdamW with stepwise learning rates (step LR), as well as combinations of $lr0 \in \{5 \times 10^{-3}, 2 \times 10^{-2}\}$, $momentum \in \{0.90, 0.98\}$, $weight\_decay \in \{1 \times 10^{-4}, 1 \times 10^{-3}\}$, $warmup \in \{0, 3\}$, $imgsz \in \{416, 640\}$, and $batch \in \{8, 24\}$. Results showed that the current configuration performed most robustly on the validation set: SGD + cosine exhibited less overfitting than AdamW. While imgsz = 640 slightly helped with very small targets, it was computationally expensive and resulted in increased convergence volatility. Batches greater than 16 were susceptible to GPU memory limitations, while batch = 8 resulted in unstable statistics and slightly reduced accuracy. To improve training efficiency and reduce video memory usage, mixed precision training (AMP) was used in the experiment, and the validation set was evaluated after each round to ensure that performance indicators and model convergence could be monitored in real time. In summary, the hyperparameters used in this paper represent a compromise between efficiency, stability, and accuracy.

Table 1. Model training parameters.

Item	Setting	Item	Setting
Model	YOLOv12x (pretrained)	Input size	512×512
Batch size	16	Epochs	500
Optimizer	SGD	Initial learning rate	1.0×10^{-2}
Final learning rate	1.0×10^{-4}	Momentum	0.937
Weight decay	5.0×10^{-4}	LR schedule	Cosine annealing
Warm-up epochs	5	Mixed precision	AMP

As shown in Figure 9, the curve trend shows that the validation loss (including bounding box regression, classification, and distribution regression) gradually stabilizes in the second half of training and reaches a minimum value of 4.712 at the 189th epoch, indicating that the model has achieved optimal performance in terms of overall error control. Furthermore, detection performance metrics remained relatively stable after reaching peaks in the middle phase: precision reached a peak of 0.757 at the 138th epoch; recall reached 0.656 at the 146th epoch, and average precision (mAP50–95) reached 0.389 at the 159th epoch. Overall, the model achieved relatively stable performance in the middle and late stages, maintaining a reasonable balance.

In terms of model selection, this study selected the weights of four key training nodes for subsequent testing to comprehensively examine the model's performance across various performance metrics. Specifically, the models from the 138th epoch (highest precision), the 146th epoch (highest recall), the 159th epoch (highest mAP50–95), and the 189th epoch (lowest validation loss) were retained and named Model A, Model B, Model C, and Model D, respectively (Table 2). This strategy not only facilitated comparisons across different metrics in subsequent experiments but also provided a diverse reference baseline for further analysis of stability and generalization performance.

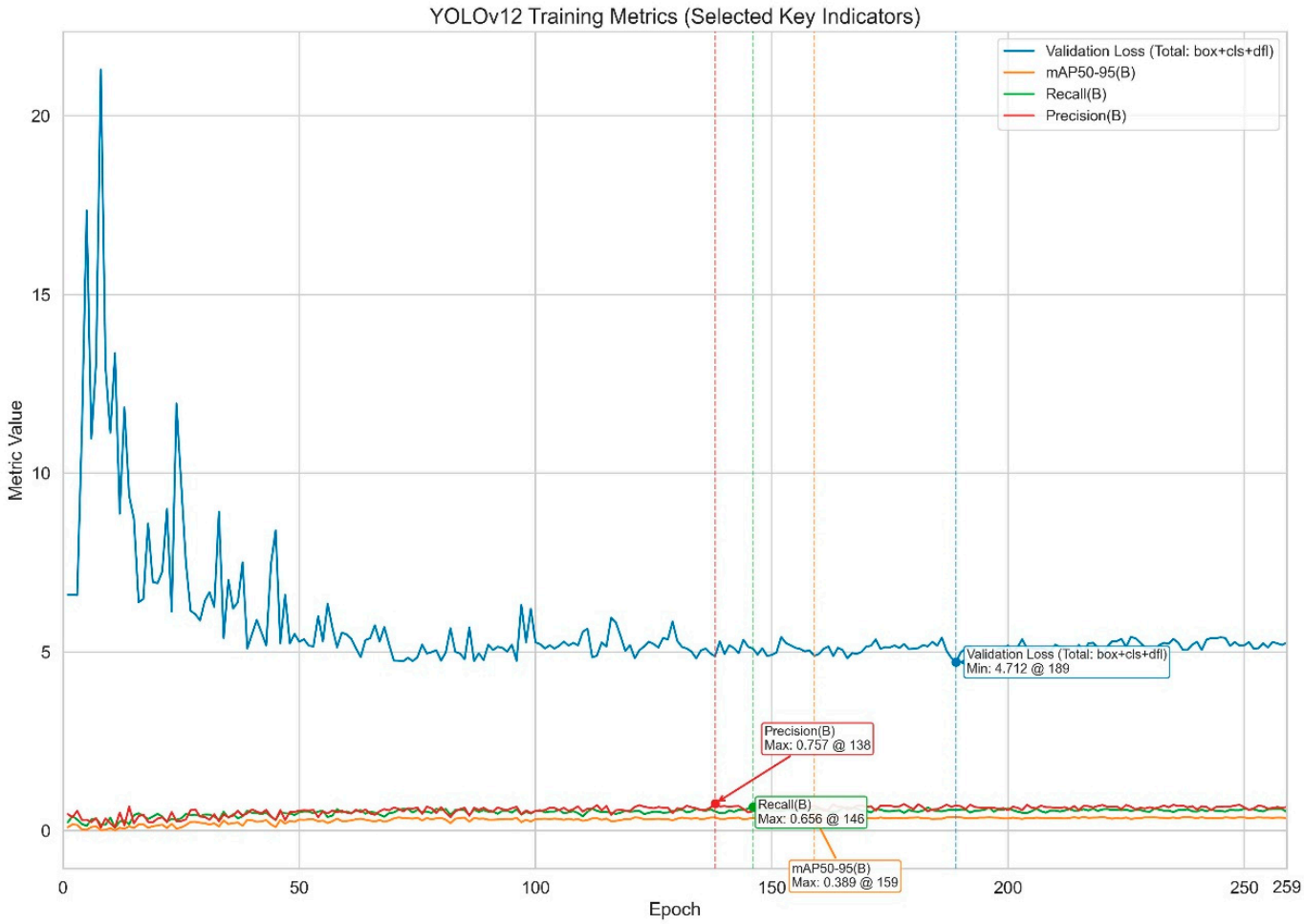


Figure 9. Numerical records during model training.

Table 2. Four key training nodes for subsequent testing.

Model Name	Four Key Training Nodes	Characteristics
A	the 138th epoch	highest precision
B	the 146th epoch	highest recall
C	the 159th epoch	highest mAP50–95
D	the 189th epoch	lowest validation loss

3.2. Model Testing Results

This study systematically evaluated the performance of four key models (A, B, C, and D) to verify their applicability to detect damage on tile surfaces. To ensure consistency across the tests, this study used a unified validation set to compare the models' performance across various metrics.

In terms of overall model performance, the metrics used include precision, recall, mean average precision (mAP@0.5 and mAP@0.5:0.95), and a comprehensive fitness metric (Fitness). As shown in Figure 10, the four models demonstrate their respective strengths and weaknesses. Model A achieves the highest precision (precision = 0.757), demonstrating its effective ability to suppress false detections. Model B excels in recall (recall = 0.655), demonstrating its broader coverage of potential impairments. Model C achieves the best performance in both mAP@0.5 (0.667) and mAP@0.5:0.95 (0.389), demonstrating its stability in both object localization and multi-threshold detection tasks. Model D, while not

achieving the highest performance in any individual metric, maintains relatively balanced performance across all metrics.

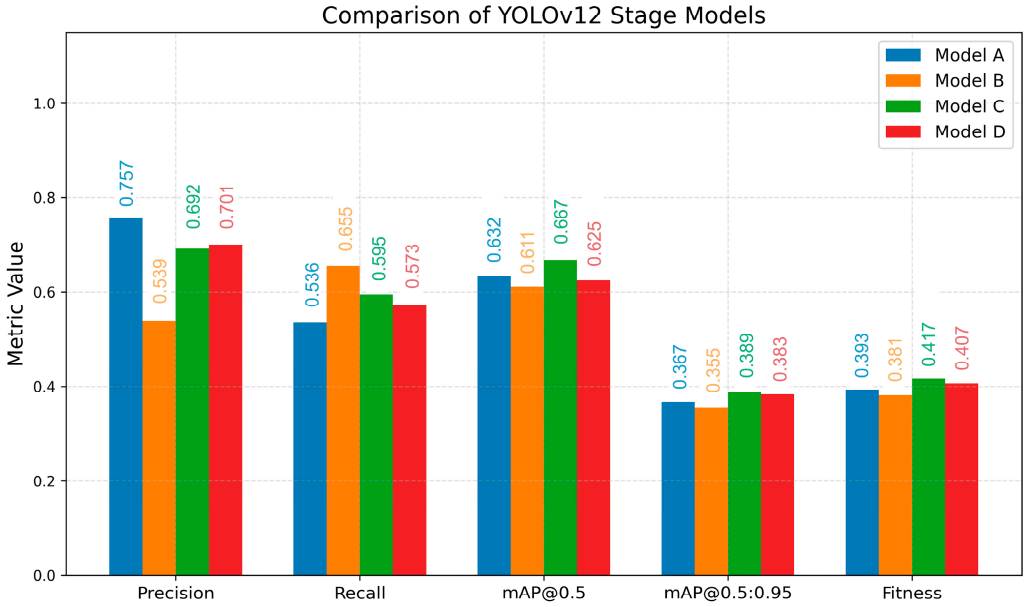


Figure 10. Comparative analysis of overall model indicators.

To further evaluate the effectiveness of different models for detecting specific damage types, this study plotted precision–recall curves for four candidate models and calculated the mean average precision (mAP) at 0.5 for each category. Compared to overall metrics, the precision–recall curve more intuitively reflects the changing trends in a model's precision and recall at different confidence thresholds, revealing the detection difficulty and adaptability of each category.

As shown in Figure 11, all four models achieved high mAP values for cracks and wear. Model C achieved a score of 0.781 for crack detection, and Model D achieved a score of 0.735 for wear detection. This suggests that the models detect damaged types with distinct linear features or clear boundaries more easily. In contrast, the detection performance for stains and water stains was generally lower, with mAP values for all four models ranging from 0.47 to 0.62. Stains, in particular, were consistently the most challenging type to detect. The difficulty may be due to their texture features being close to the background, their morphology being highly variable, and their annotated boundaries being blurred. Overall, Model C achieved an average mAP@0.5 score of 0.667, maintaining the best overall performance among the four models. Models A and D achieved relatively similar overall performances, while Model B maintained a relatively balanced performance across all categories.

Furthermore, the confusion matrix reflects the proportion of each type of damage that is correctly identified or misclassified as another category during the prediction process, revealing the strengths and weaknesses of the model in multi-class detection tasks. As shown in Figure 12, the diagonal values indicate that cracks and wear are generally well identified, while stains and water stains still exhibit significant confusion. For example, Model A achieves an accuracy of 0.73 for cracking and 0.76 for wear, but only 0.67 for stains, with 0.32% misclassified as background. Model B improves its performance for cracking to 0.78. Model C demonstrates the most balanced overall performance across multiple categories, achieving 0.73 and 0.77 for cracking and wear, respectively, and maintaining 0.60 and 0.65 for water stain and stains. In contrast, Model D achieves the lowest accuracy for stains, at only 0.51, and the highest misclassification rate for background (0.47), demonstrating its

limited robustness against complex background interference. A further comparison reveals that background classification is a key factor influencing overall performance. Consistent with the PR curve results in the previous section, stains and water stain remain the most challenging categories; while cracking and wear maintain high consistency and stability across all models.

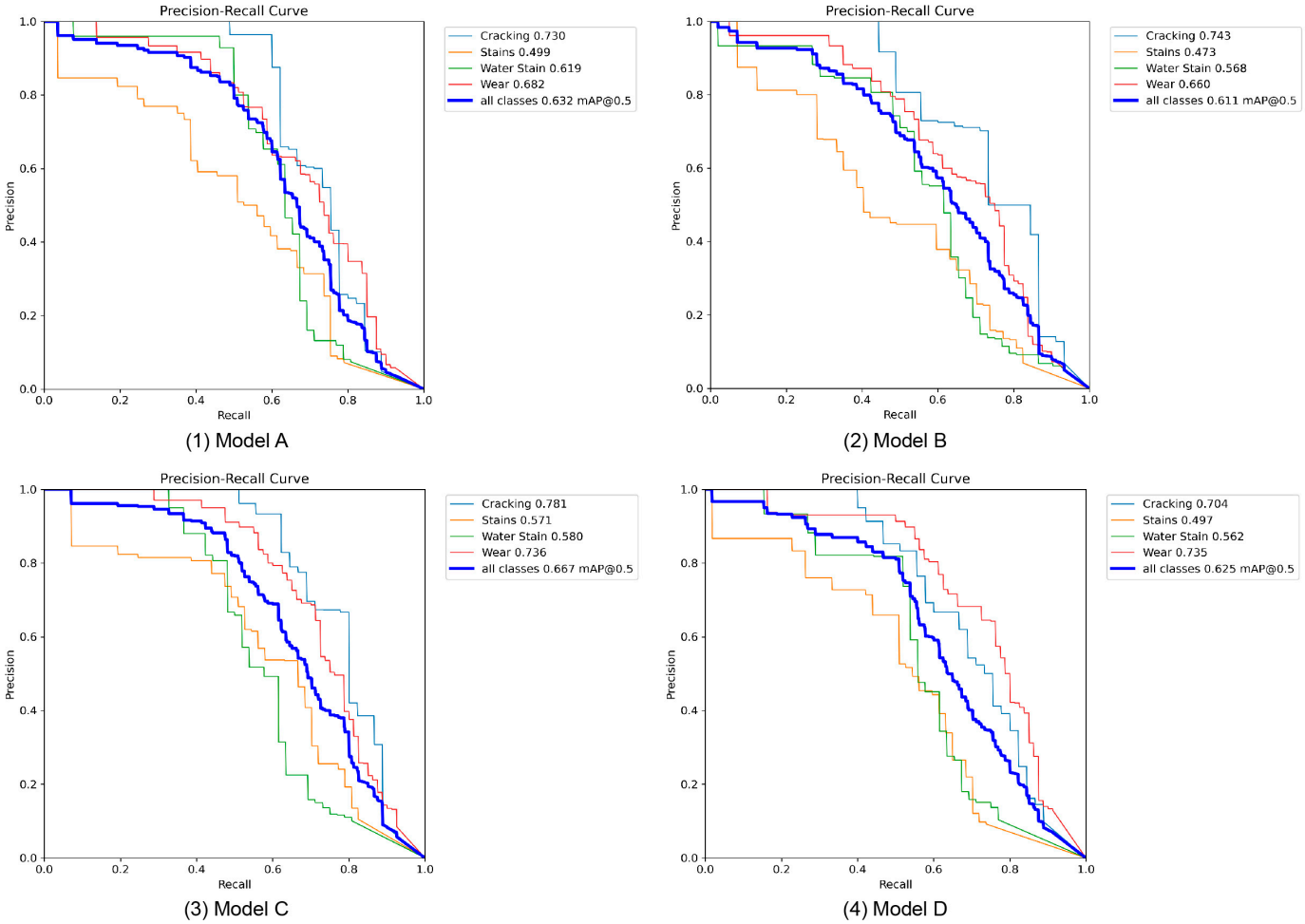


Figure 11. Comparative analysis of model precision–recall curves.

To verify the model's fit on the training data, this study selected a portion of the training set samples for testing, with the results shown in Figure 13. By comparing the prediction performance of the four candidate models on the same image, we can further observe the model's recognition accuracy and confidence distribution for different damage categories.

From the overall performance perspective, all four models can accurately identify target damage in the training set, but there are differences in performance in different categories and scenarios. In the sample of Figure 13(1), all four models can detect multiple water stains, but the number and position of the prediction boxes of Models A and C are more consistent with the actual damage area, with confidence levels in the range of 0.44–0.81, while Models B and D have some box overlaps or low confidence levels. Figure 13(2) shows a combined scene of cracks and stains. All four models correctly identify the main cracks, and the confidence levels are all above 0.70. However, in the detection of stains, the predictions of Models C and D are more stable (confidence levels 0.77 and 0.84), while the detection box boundaries of Models A and B are not accurate enough. The results in Figure 13(3,4) further reflect the differences between the models: in terms of crack detection, all four models can output stable results, but in the identification

of water stains and stains, the confidence levels of Models C and D are generally higher than those of Models A and B, and the number of false detections is smaller.

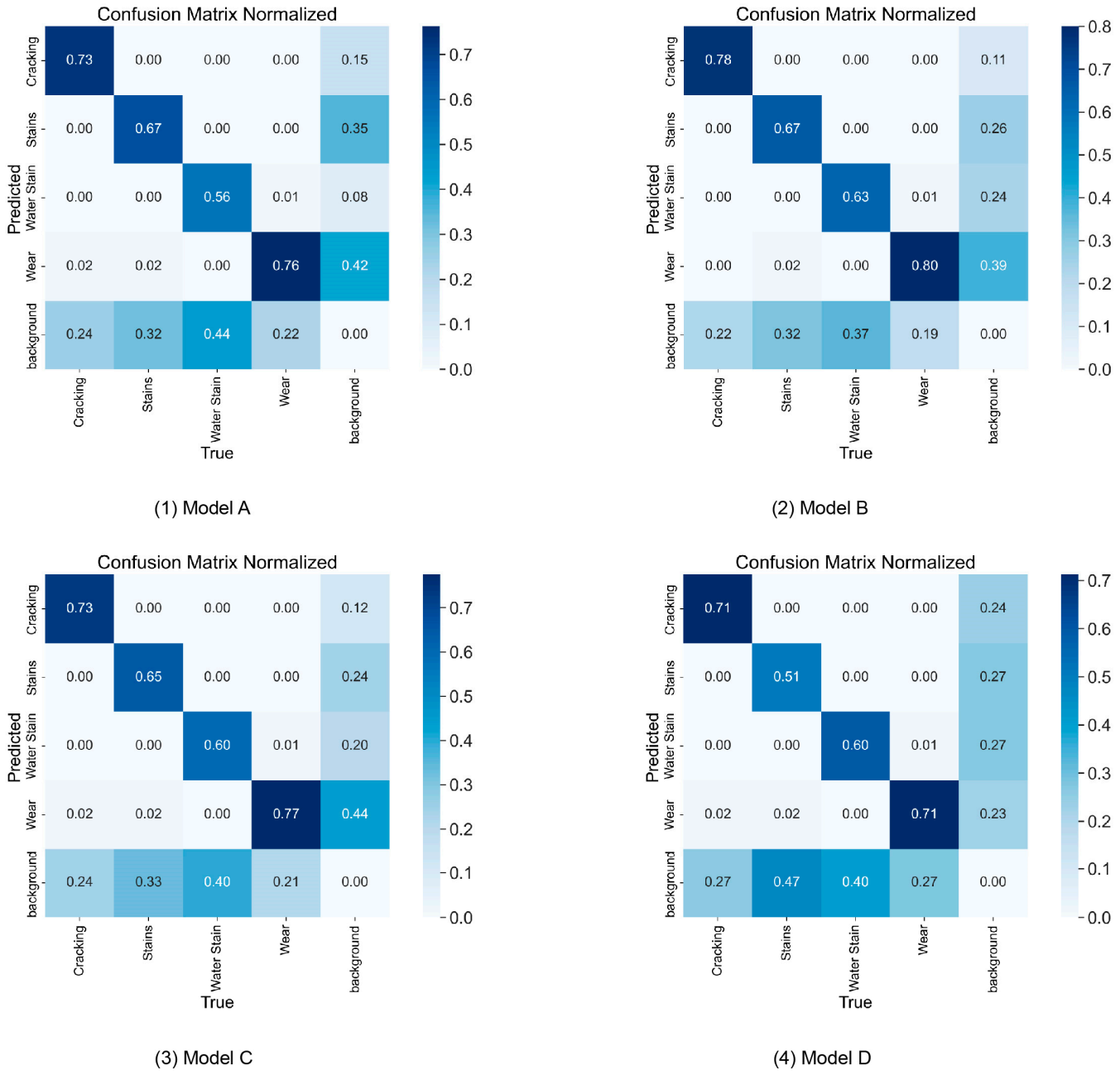


Figure 12. Model confusion matrix comparison analysis.

To further verify the generalization ability of the model in real scenes, this study used the original images newly collected on-site (4284 × 5712 pixels, straight out of the camera) for testing, and the results are shown in Figure 14. In Figure 14(1), Model A only identified the crack in the upper-right corner, with the least detection. Model B can identify the cracks in the middle and upper parts but misses the crack in the lower-left corner, and water stain recognition is inaccurate, and the water stain in the upper-left corner is not detected. Model C is similar to Model B and also has the above omissions; Model D also has missed detections but can recognize the water stain in the upper-left corner and detect the crack in the lower-left corner. In Figure 14(2), the results of Model B are relatively accurate. Although the crack in the lower-right corner is missed, the rest of the stains

and cracks are correctly identified; Model A and Model D mistakenly identify the toes of the photographer on the left side of the picture as cracks; Model C can identify the main damage but misses the stain in the upper-right corner and can detect the crack in the lower-right corner. Figure 14(3) shows that Model C has the most accurate results, achieves the least missed detections, and basically covers all damage types. In comparison, Models A, B, and D all have different degrees of missed detections, with Models A and B having the most obvious missed detections. In Figure 14(4), only Model C can accurately identify five water stains in the image, while the other models have different degrees of missed detections. It should be noted that Model C also missed a crack mark, and Model D also misidentified an area as wear.

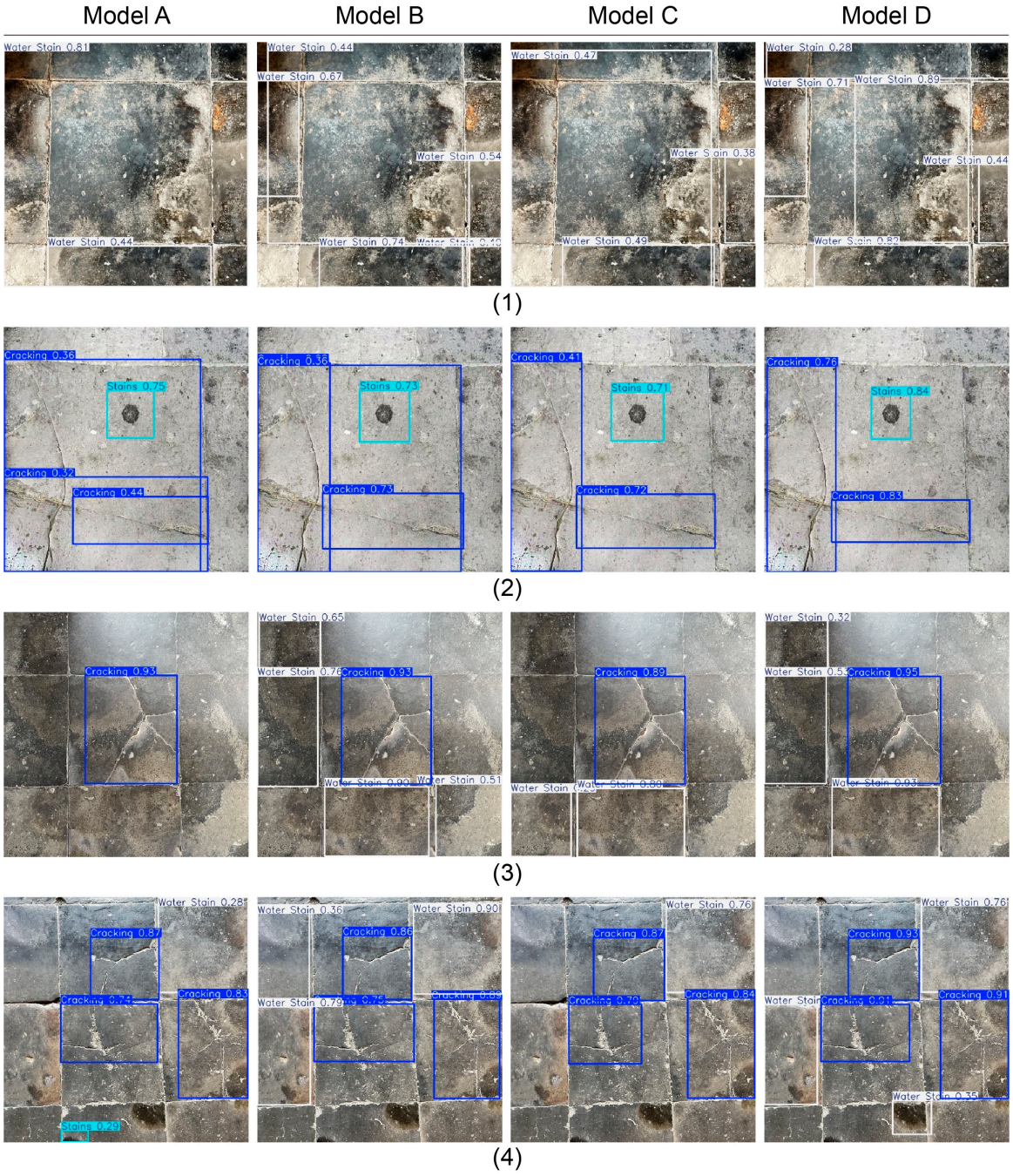


Figure 13. The test results of the model in the training set. (**1**–**4**) represent the detection results of the model for different reduction-fired cyan square bricks.

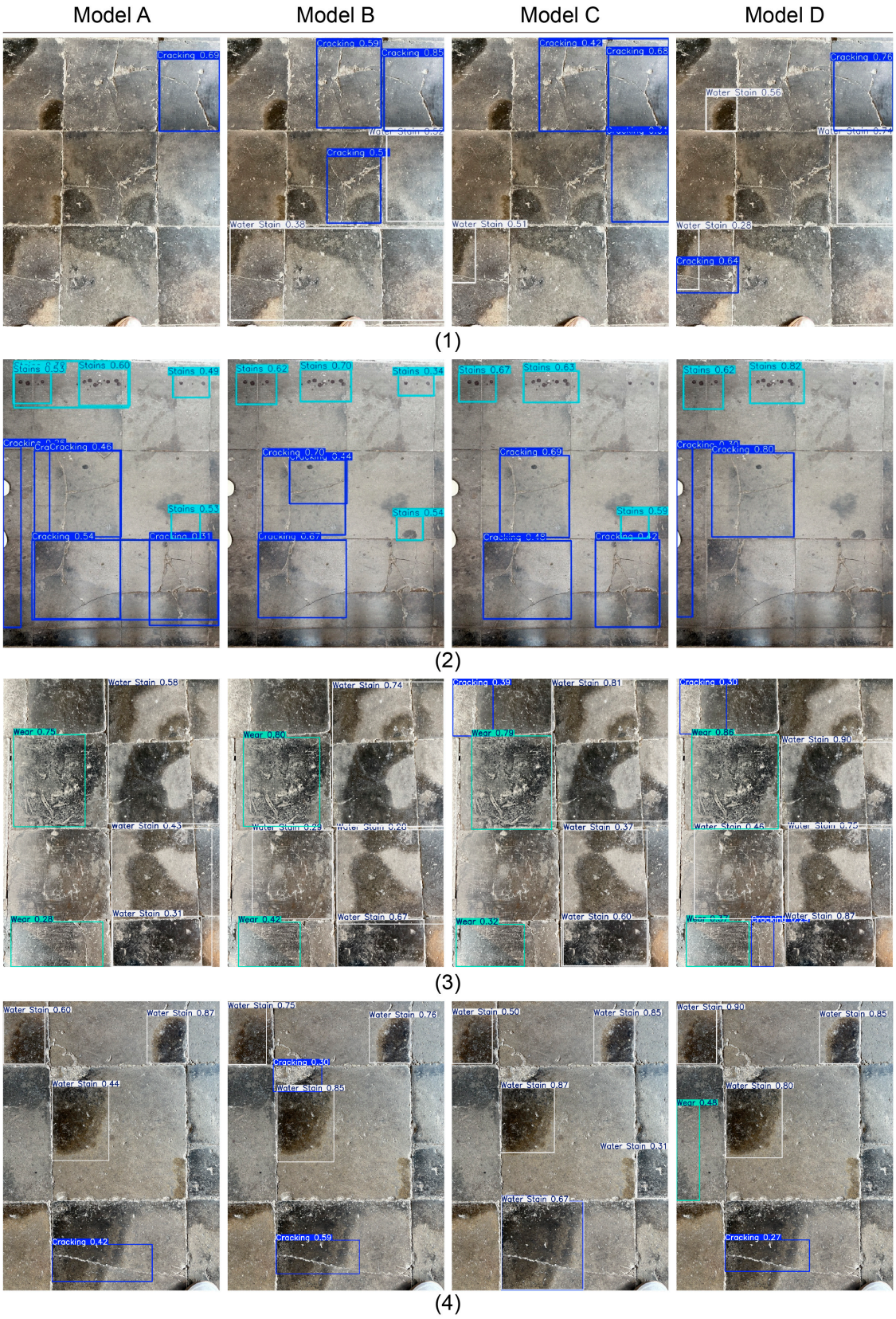


Figure 14. Test results of the model in real scenarios. (**1**–**4**) represent the detection results of the model for different reduction-fired cyan square bricks.

In summary, this study systematically evaluated the performance of four candidate models (A, B, C, and D) in tile surface damage detection. The results indicated that

Model A achieved the highest precision (precision = 0.757), Model B achieved the best recall (recall = 0.691), Model D offered balanced performance but suffered from insufficient robustness, and Model C led in comprehensive metrics (mAP@0.5 = 0.667), maintaining stability in multi-class detection and demonstrating the strongest generalization ability in complex backgrounds. Combining overall metrics, class discrimination, training fit, and real-world testing, Model C outperformed the other models in accuracy, stability, and adaptability. Therefore, it is suitable for further application and provides reliable support for automated floor damage detection.

4. Discussion

Jiangnan gardens, a vital part of world heritage, feature paving made of reduction-fired cyan square bricks (RCSBs). These bricks not only embody the garden's spatial layout and cultural heritage but are also frequently used by visitors and exposed to the elements. Long-term wind and rain erosion, along with heavy foot traffic, commonly cause wear, cracking, water stains, and other damage to the tiles. If neglected, these damages not only jeopardize the floor's esthetics and safety but also hasten the deterioration of the heritage material. Therefore, developing an intelligent, non-contact, and quantifiable RCSB damage detection method has become a crucial task in cultural heritage preservation. Based on the YOLOv12 model, this study attempts to construct a comprehensive process from image input and feature extraction to damage segmentation and statistics, providing a technical path for the monitoring and preventative protection of RCSBs in gardens.

4.1. Application in Actual Scenarios

As shown in Figure 14, in the test experiment (Figure 15(1–4)), this study used the same 512 × 512-pixel samples in the training phase to perform inference detection on representative damage types. The results show that the model can stably identify multiple types of damage. In the cracked sample in Figure 15(1), the thermal map shows that the backbone, neck, and head of the model all produced concentrated responses in the crack area. The segmentation module effectively outlined the crack boundary, and the statistical module further calculated the number and area of the cracks. In the water stain and stain samples in Figure 15(2), the model can distinguish between local dark water stains and the background. In the complex surface sample in Figure 15(3), the model can still lock the main crack and pollution areas under conditions of more interference, showing a certain degree of robustness. In Figure 15(4), the model identified a large area of water stains, verifying its applicability in heavily polluted environments. Overall, the thermal response and the segmentation results show a favorable correspondence, indicating that feature aggregation at each level of the network can effectively support the final recognition and counting.

Furthermore, this study's method not only provides typological identification of reduction-fired cyan square brick damage but also achieves quantitative output through segmentation and counting, which has direct implications for project management. Compared to traditional methods that rely on manual visual inspection, AI-based inspection can significantly improve efficiency in large-scale garden reduction-fired cyan square brick inspections and reduce subjective bias in human judgment. More importantly, heatmap visualization provides diagnostic evidence for subsequent research, aiding in understanding the model's focusing mechanism and serving as a reference for manual inspections.

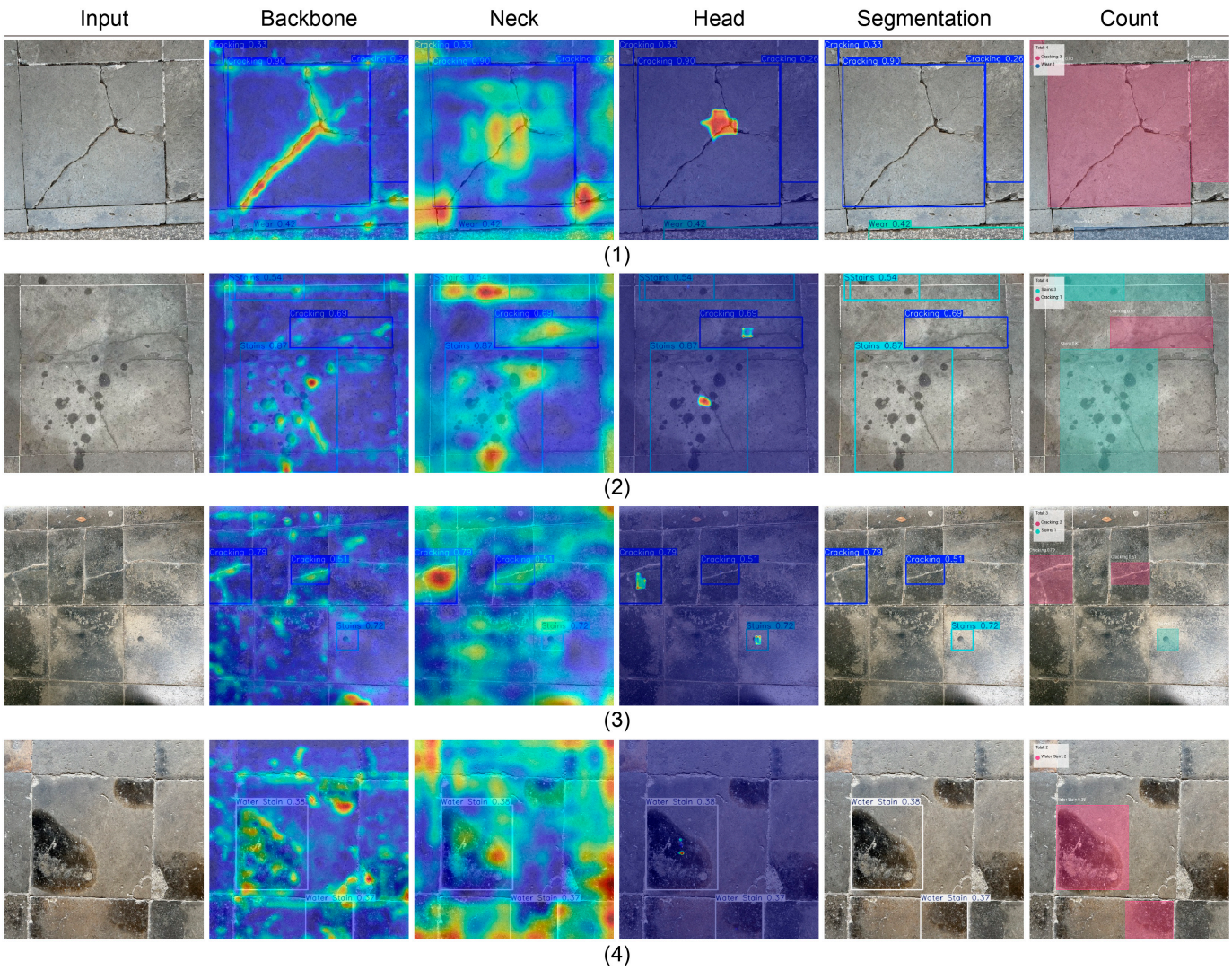


Figure 15. The detection process and results of the model on the ground of Jiangnan gardens. (**1**–**4**) represent the detection results of the model for different reduction-fired cyan square bricks.

4.2. Design of Ground Intelligent Detection Equipment

In response to the characteristics and protection needs of Jiangnan classical garden's reduction-fired cyan square bricks, this study designed an intelligent ground detection application model. The trained algorithm model can be embedded in an intelligent ground detection terminal with cleaning capabilities [37], creating an intelligent detection device that can not only identify ground damage and but also provide cleaning capabilities (Figure 16).

In terms of inspection capabilities, the intelligent inspection equipment is equipped with a high-resolution camera and an infrared thermal imaging module. The camera can capture visual damage such as wear, cracks, water stains, and stains. Furthermore, it can be equipped with LiDAR technology to construct a three-dimensional model of surface damage on reduction-fired cyan square bricks, accurately measuring the size and depth of the damaged area and providing a quantitative basis for subsequent repairs.

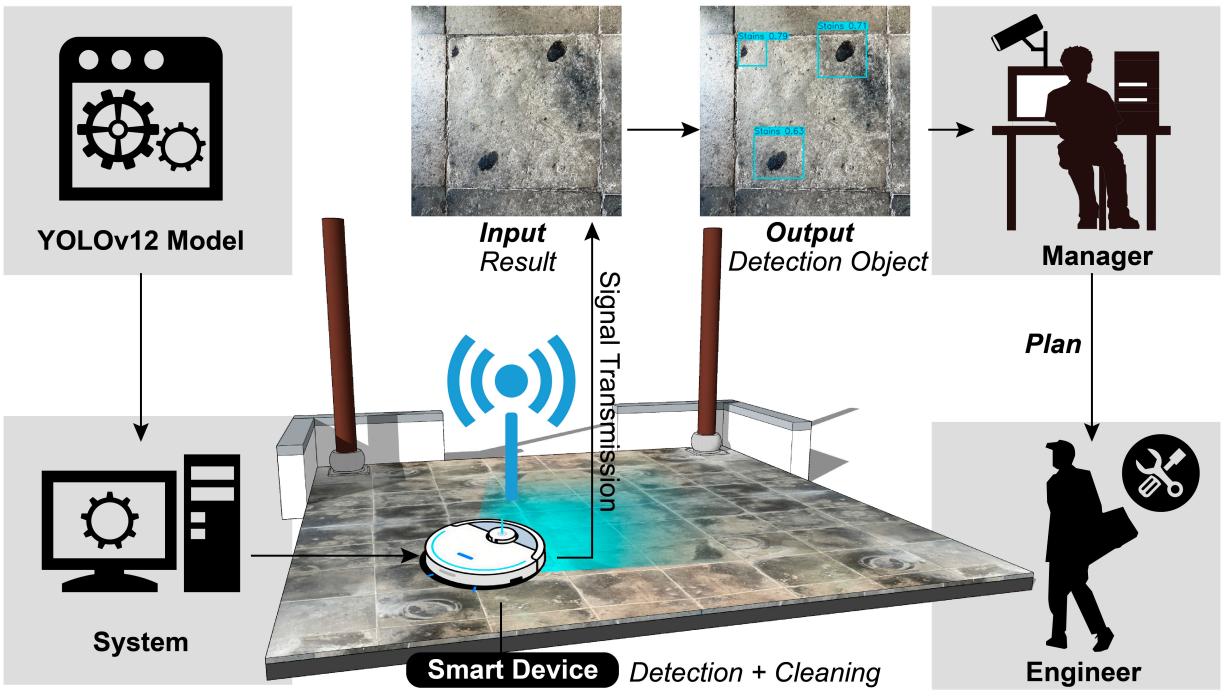


Figure 16. Design strategy for the ground intelligent detection application model.

In actual use, the device is designed with dual modes of local storage and cloud synchronization. For example, during operation, the device saves inspection data and cleaning records in real time. Furthermore, when network conditions permit, the data is uploaded to the cloud platform for landscape management in real time. Managers can log in to the platform via a computer or mobile device to view inspection reports and cleaning status, clearly understanding the overall condition and cleanliness of the reduction-fired cyan square bricks. The platform automatically issues an alert when it detects severe damage, prompting managers to promptly address it. Repair personnel can assign suitable materials for localized surface repairs to address minor wear issues. Severe cracking and extensive damage require replacements. The 3D model data provided by the device enables craftsmen to precisely prefabricate replacement tiles, reducing on-site work time. After the repair is completed, the intelligent inspection equipment will focus on monitoring the repaired area during the next inspection to assess the quality and ensure that the repair meets requirements and the surface is clean.

Researchers believe that the integration of artificial intelligence and intelligent ground-based inspection equipment will provide an efficient and sustainable approach for monitoring and protecting the reduction-fired cyan square bricks in Jiangnan gardens. This closed-loop process of "cleaning-inspection-repair-re-monitoring" can promptly detect and address damage to the reduction-fired cyan square bricks in Jiangnan classical gardens while maintaining the pristine surface of the bricks. This technology-driven inspection model is significant to world heritage and provides garden managers and artisans with a foundation for restoration and convenient inspection tools.

5. Conclusions

5.1. Research Discovery

This study aims to address the problem of surface damage that is prone to occur in the reduction-fired cyan square brick pavement in the Jiangnan garden world heritage scene under the dual effects of long-term humidity, heat, and high traffic, and it proposes a non-contact, automated intelligent detection solution. The purpose of this study is to

achieve robust recognition of four types of damage, namely cracking, stains, water stains, and wear, while maintaining good generalization under complex textures and low-contrast backgrounds. In terms of methodology, a multi-scene dataset was constructed, and end-to-end training was performed at a scale of 512 × 512 using YOLOv12 as the backbone. Four types of models were selected according to key training nodes for comparison and tested on multiple sets of data. The main conclusions of this study are as follows: (1) In terms of the overall indicators of the validation set, Model C achieved the best comprehensive accuracy (mAP@0.5 = 0.667), taking into account both precision and recall. (2) Analysis of the PR and confusion matrix by category showed that cracking and wear were easier to distinguish, while stains and water stains were generally low (mostly between 0.47 and 0.62) due to the influence of boundary blur and texture similarity. Model C had a more balanced diagonal accuracy in the confusion matrix and significantly lower background confusion. (3) In the training set samples and the original site images, Model C had the lowest missed detection rate and remained stable in the scene with multiple damage types, reflecting the best generalization ability and engineering usability. At the same time, the shortcomings of this study are as follows: the data scale and category are unbalanced; the recall of stains and water stains under extreme lighting and shadows is still limited, and currently only two-dimensional representation information is provided.

5.2. Limitations and Future Work

However, the limitations of this study still need to be viewed objectively: (1) Although the model performs well in typical samples, it may still produce misjudgments or missed detections under extreme lighting conditions, complex shadows, or severely weathered reduction-fired cyan square brick surfaces. (2) The scale and diversity of training data are still limited. In the future, more garden reduction-fired cyan square brick samples from different environments and ages need to be introduced to enhance the generalization ability of the model. (3) The output of the model is currently mainly concentrated at the two-dimensional image level. It is not yet possible to fully characterize three-dimensional features such as the depth of cracks and the thickness of wear, limiting its application value in structural safety assessments. (4) Although YOLOv12 offers good detection accuracy and speed, it is necessary to further explore its potential for integration with drone images, three-dimensional scanning, and environmental monitoring data within the long-term monitoring system of heritage protection to achieve a more comprehensive diagnosis and early warning.

Future research can be carried out in the following directions: (1) Introducing multi-scale feature fusion and domain adaptive training strategies: In terms of multi-scale feature fusion, by constructing a fusion network containing shallow high-resolution features and deep semantic features, the model can not only capture the detailed information of the wear texture and tiny cracks on the surface of the tiles but also understand the macro-damage characteristics such as large-area water stains and patchy stains, thereby improving the recognition ability of damage of different sizes. (2) Implementing a category reweighting mechanism: Considering that the number of samples of different damage types in the training data may be unbalanced—for example, the number of common stain samples is far greater than the number of rare long and thin crack samples—by dynamically calculating the weights of each category of samples, a higher weight is given to damage types with a small number of samples and high recognition difficulty, and attention is paid to these categories during model training. (3) Constructing a multimodal (infrared/humidity) and three-dimensional data acquisition fusion framework: In terms of multimodal data acquisition, in addition to conventional visible light images, the introduction of infrared images can capture the temperature differences caused by damage inside the reduction-fired cyan

square bricks. Humidity data can reflect the degree and range of water stain penetration, providing more dimensional information for damage recognition. Three-dimensional data acquisition uses laser scanning and other technologies to obtain three-dimensional point cloud data of the floor tile surface, accurately characterizing three-dimensional morphological parameters such as crack depth and wear thickness. (4) Improvement of the loss function and optimization of the post-processing algorithm for slender crack detection: The model often overlooks slender cracks because of their narrow shape and unclear features. In the improvement of the loss function, a loss term specifically targeting slender targets is designed to increase the focus on crack edges and continuity, guiding the model to more accurately locate and identify slender cracks. Through the above technical path, the recall rate and precision balance performance of the model can be further improved, providing more solid technical support for the long-term and normalized monitoring of heritage protection scenarios.

Author Contributions: Conceptualization, L.Y. and Y.C.; methodology, L.Y. and L.Z.; software, L.Y.; validation, L.Y. and L.Z.; formal analysis, L.Y., Y.C., and L.Z.; investigation, L.Y., and X.J.; resources, L.Y., and X.J.; data curation, L.Y.; writing—original draft preparation, L.Y.; writing—review and editing, L.Y. and Y.C.; visualization, L.Y. and X.J.; supervision, Y.C.; project administration, Y.C.; funding acquisition, Y.C. and L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Faculty Research Grants funded by (1) Huangshan City Social Sciences Innovation and Development Research Project (grant number: 2025297); (2) Macau University of Science and Technology (grant number: FRG-25-041-FA; FRG-25-067-FA); and (3) Guangdong Provincial Department of Education's key scientific research platforms and projects for general universities in 2023: Guangdong, Hong Kong, and Macau Cultural Heritage Protection and Innovation Design Team (grant number: 2023WCXTD042). The funders had no role in study conceptualization, data curation, formal analysis, the methodology, software, the decision to publish, or the preparation of the manuscript. There was no additional external funding received for this study.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Lina Yan organized and participated in this study's investigation and has all the original data. To anyone who is interested or wants to learn more, please contact Lina Yan (linayan@shzq.edu.cn) for further information.

Conflicts of Interest: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Appendix A. Experimental Environment

Table A1. Experimental environment.

Category	Version	Category	Version
Operating System	Windows 10 (22H2) Professional Edition/AMD64 Architecture	Python	3.11.13
Processor	AMD Ryzen 7 9800X3D 8-Core Processor	System memory	64 GB
Pytorch	2.7.0 + cu128	Torch cuda	12.8
GPU	NVIDIA GeForce RTX 5070 Ti 16 GB	CUDA	12.9.41
Flash attn	2.7.4.post1	NumPy	1.26.4
Pandas	2.3.0	Matplotlib	3.10.3
OpenCV	4.9.0	Ultralytics	8.3.63

Appendix B. Model Structure

Table A2. Model Structure.

Stage	Layer/Module	Output Channels	Stride	Notes
Backbone	Conv	96	2	Stem conv
	Conv (groups = 2)	192	2	Downsample
	C3k2 (2 × C3k)	384	1	Bottleneck blocks
	Conv (groups = 4)	384	2	Downsample
	C3k2 (2 × C3k)	768	1	Bottleneck blocks
	Conv	768	2	Downsample
	A2C2f (4 × ABlock)	768	1	Attention-enhanced
	Conv	768	2	Downsample
	A2C2f (4 × ABlock)	768	1	Transformer-like
Neck	Upsample + Concat	-	-	Feature fusion
	A2C2f (2 × C3k)	768	1	Fusion block
	Upsample + Concat	-	-	PANet-like fusion
	A2C2f (2 × C3k)	384	1	Fusion block
	Conv	384	2	Downsample
	Concat + A2C2f (2 × C3k)	768	1	Fusion block
	Conv	768	2	Downsample
	Concat + C3k2 (2 × C3k)	768	1	Fusion block
Head	Detect	-	-	3 detection scales, with DFL

References

1. Jiang, J.; Zhou, T.; Han, Y.; Ikebe, K. Urban Heritage Conservation and Modern Urban Development from the Perspective of the Historic Urban Landscape Approach: A Case Study of Suzhou. *Land* **2022**, *11*, 1251. [CrossRef]
2. Jiang, J.; Zang, T.; Xing, J.; Ikebe, K. Spatial Distribution of Urban Heritage and Landscape Approach to Urban Contextual Continuity: The Case of Suzhou. *Land* **2023**, *12*, 150. [CrossRef]
3. Dai, X. The Chinese City Suzhou in Seven Hundred Years: An Investigation of the Relationship Between the Changing Functional Pattern and its Spatial Structure in the Urban Transformation Process. Master's Thesis, UCL (University College London), London, UK, 2004. Available online: https://discovery.ucl.ac.uk/id/eprint/1431 (accessed on 19 August 2025).
4. Wang, L.; Shen, J.; Chung, C.K.L. City profile: Suzhou-a Chinese city under transformation. *Cities* **2015**, *44*, 60–72. [CrossRef]
5. Huang, Y.; Meadows, M.E. Challenges to the Sustainability of Urban Cultural Heritage in the Anthropocene: The Case of Suzhou, Yangtze River Delta, China. *Land* **2025**, *14*, 778. [CrossRef]
6. Classical Gardens of Suzhou. Available online: https://whc.unesco.org/en/list/813/ (accessed on 19 August 2025).
7. "The Best Brick in the World": Suzhou Imperial Kiln Gold Brick. Available online: https://www.js-fz.cn/n96/20211111/i9718.html (accessed on 19 August 2025).
8. Xia, W. An Investigation of Chinese Historical Grey Bricks of Soochow, Jiangsu and the Effect of Tung Oil Treatment. Master's Thesis, University of Pennsylvania, Philadelphia, PA, USA, 2015. Available online: https://core.ac.uk/download/pdf/76388462.pdf (accessed on 19 August 2025).
9. Dunzhen, L.; Wood, F. The Traditional Gardens of Suzhou ("Su zhou gu dian yuan lin"). *Gard. Hist.* **1982**, *10*, 108–141. [CrossRef]
10. Koss, S.L. Suzhou: A history of transformations. In *Suzhou in Transition*; Routledge: Abingdon, UK, 2020; pp. 15–41.
11. Wu, Y.; Wu, Y.; Mao, Y. Check for updates Research on the Transformation and Development of Modern Western-Style Villas in Suzhou Under the Concept of Sustainable Development. In Proceedings of the 2022 International Conference on Green Building, Civil Engineering and Smart City, Guilin, China, 8–10 April 2022; Springer Nature: Singapore, 2022; Volume 211, p. 190. [CrossRef]
12. Wu, Y.; Yonglan, W.; Mingrui, M.; Muxian, W.; Li, X.; Candeias, A. Research on Construction of Suzhou's Historical Architectural Heritage Corridors and Cultural Relics-Themed Trails Based on Current Effective Conductance (CEC) Model. *Buildings* **2025**, *15*, 2605. [CrossRef]
13. Yonghui, L.; Xie, H.; Wang, J.; Li, X. Experimental study of the isothermal sorption properties of late Qing and 1980s grey bricks in Wujiang, Suzhou, China. *Front. Archit. Res.* **2013**, *2*, 483–487. [CrossRef]
14. Wang, S.; Gao, J.; Wang, X.; Wu, D.; Pan, Y.; Xu, M. Historical and Physicochemical Analysis of the Clinker Bricks in the Smart Memorial Gymnasium of the Tiancizhuang Campus at Soochow University, China. *Buildings* **2023**, *13*, 161. [CrossRef]

15. Chun, Q.; Dong, Y.; van Balen, K.; Xu, X. Experimental research on material properties of ancient white bricks in the Yichun Region, China. *Int. J. Archit. Herit.* **2017**, *11*, 554–565. [CrossRef]
16. Li, Y.; Ma, Y.; Xie, H.; Li, J.; Li, X. Cross validation of hygrothermal properties of historical Chinese blue bricks with isothermal sorption experiments. *Front. Archit. Res.* **2021**, *10*, 164–175. [CrossRef]
17. Zhao, P.; Zhang, X.; Qin, L.; Zhang, Y.; Zhou, L. Conservation of disappearing traditional manufacturing process for Chinese grey brick: Field survey and laboratory study. *Constr. Build. Mater.* **2019**, *212*, 531–540. [CrossRef]
18. Zhang, Y.; Chen, Z.; Yue, J.; Xu, S. Experimental study on improving the performance of historical grey bricks by amide polymer. *Constr. Build. Mater.* **2025**, *489*, 142238. [CrossRef]
19. Yan, L.; Chen, Y.; Zheng, L.; Zhang, Y. Application of computer vision technology in surface damage detection and analysis of shedthin tiles in China: A case study of the classical gardens of Suzhou. *Herit. Sci.* **2024**, *12*, 72. [CrossRef]
20. Chen, X.; He, J.; Wang, S. Deep learning-driven pathology detection and analysis in historic masonry buildings of Suzhou. *npj Herit. Sci.* **2025**, *13*, 197. [CrossRef]
21. YOLO12: Attention-Centered Object Detection. Available online: https://docs.ultralytics.com/zh/models/yolo12/ (accessed on 19 August 2025).
22. Tian, Y.; Ye, Q.; Doermann, D. Yolov12: Attention-centric real-time object detectors. *arXiv* **2025**, arXiv:2502.12524. [CrossRef]
23. Sapkota, R.; Qureshi, R.; Calero, M.F.; Badjugar, C.; Nepal, U.; Poulose, A.; Karkee, M. Yolov12 to its genesis: A decadal and comprehensive review of the you only look once (yolo) series. *arXiv* **2024**, arXiv:2406.19407. [CrossRef]
24. Lee, W.S. Status of the Materials Used in Suzhou Gardens in the Late Ming Dynasty. Master's Thesis, University of Hawai'i at Manoa, Honolulu, HI, USA, 2006.
25. Cao, T.; Wang, X. Investigation of New Private Gardens in Yangtze River Delta and Reflection on Design Value. In *Design, User Experience, and Usability: Design for Emotion, Well-Being and Health, Learning, and Culture, Proceedings of the 11th International Conference, DUXU 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, 26 June–1 July 2022*; Soares, M.M., Rosenzweig, E., Marcus, A., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2022; Volume 13322. [CrossRef]
26. Liu, K.; Wang, Y.; Yang, R.; Xian, Z.; Takeda, S.; Zhang, J.; Xing, S. Interpreting the space characteristics of everyday heritage gardens of Suzhou, China, through a space syntax approach. *J. Asian Archit. Build. Eng.* **2024**, *24*, 3471–3496. [CrossRef]
27. Cao, Y.; Yuan, Q.; Bi, S.; Qin, L.; Yu, K. Research on the regenerative design and practice of the residential architecture atlases in Jiangnan during the Ming and Qing dynasties: A case study of the Mingyue Bay village in Suzhou. *J. Asian Archit. Build. Eng.* **2024**, *24*, 3381–3402. [CrossRef]
28. Feng, Y.; Gao, C.; Tong, X.; Chen, S.; Lei, Z.; Wang, J. Spatial Patterns of Land Surface Temperature and Their Influencing Factors: A Case Study in Suzhou, China. *Remote Sens.* **2019**, *11*, 182. [CrossRef]
29. Chen, Y.; Zhang, N.; Zhu, Y. High-Resolution Simulations of the Urban Thermal Climate in Suzhou City, China. *Atmosphere* **2019**, *10*, 118. [CrossRef]
30. Wang, Y.; Abuel-Naga, H. Unfired Bricks from Wastes: A Review of Stabiliser Technologies, Performance Metrics, and Circular Economy Pathways. *Buildings* **2025**, *15*, 1861. [CrossRef]
31. Wood, N. *Chinese Glazes: Their Origins, Chemistry, and Recreation*; University of Pennsylvania Press: Philadelphia, PA, USA, 1999.
32. Valenstein, S.G. *A Handbook of Chinese Ceramics*; Metropolitan Museum of Art: New York, NY, USA, 1989.
33. Xu, Y.; Xu, Y.; Wang, Y.; Wu, L.; Li, G.; Song, S. Spatial and temporal trends of reference crop evapotranspiration and its influential variables in Yangtze River Delta, eastern China. *Theor. Appl. Climatol.* **2017**, *130*, 945–958. [CrossRef]
34. Chen, S.; Su, Y.; Fang, X.; He, J. Climate records in ancient Chinese diaries and their application in historical climate reconstruction–a case study of Yunshan Diary. *Clim. Past Discuss.* **2020**, *16*, 1873–1887. [CrossRef]
35. Alif, M.A.R.; Hussain, M. Yolov12: A breakdown of the key architectural features. *arXiv* **2025**, arXiv:2502.14740. [CrossRef]
36. Khanam, R.; Hussain, M. A Review of YOLOv12: Attention-Based Enhancements vs. Previous Versions. *arXiv* **2025**, arXiv:2504.11995. [CrossRef]
37. Zheng, L.; Chen, Y.; Yan, L.; Zhang, Y. Automatic Detection and Recognition Method of Chinese Clay Tiles Based on YOLOv4: A Case Study in Macau. *Int. J. Archit. Herit.* **2023**, *18*, 1551–1570. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.