

A Deep Learning Approach to Intelligent Detection of Shedthin Tile Pathology in Suzhou Classical Gardens

Xi Chen ¹, Jinwen Wang ², Shiruo Wang ^{3*}

¹ School of Architecture, Soochow University, 199 Ren'ai road, Suzhou Industrial Park, Suzhou, China - xchen1@suda.edu.cn

² School of Architecture, Soochow University, 199 Ren'ai road, Suzhou Industrial Park, Suzhou, China - 20244241032@stu.suda.edu.cn

^{3*} School of Architecture, Soochow University, 199 Ren'ai road, Suzhou Industrial Park, Suzhou, China – srwang@suda.edu.cn

Keywords: Deep Learning, Architectural heritage Conservation, YOLO11-seg, Suzhou Classical Gardens, Shedthin tiles.

Abstract

Under the current rapid development of digitalization and artificial intelligence in heritage practice, deep-learning-driven pathology detection is emerging as a pivotal tool for preventive conservation. Deep learning-based intelligent pathology detection in heritage conservation has garnered increasing attention. This study explores intelligent pathology detection techniques using the YOLO11-seg model, taking the pathology identification of shedthin tiles of Suzhou classical gardens as a case study. Through data collection and annotation of 1,250 high-resolution images of shedthin tiles, a training dataset was constructed. After 362 training epochs, the model achieved automatic recognition of four key pathological type-water stains, color aberration, surface scaling, and excessive gaps-with respective accuracies of 79.31%, 73.38%, 61.12%, and 75.60%, and an overall accuracy of 74.38% that meets practical application requirements.that generally meets practical application requirements. The study further conducted quantitative analysis of detection results to assess the severity of shedthin tiles damage, providing critical references for formulating scientific restoration strategies. Compared with traditional visual surveys, the proposed workflow (i) increases detection speed by an order of magnitude, (ii) standardises assessments across inspectors, and (iii) captures early-stage micro-pathologies often overlooked in manual inspections. The results demonstrate that how integrating deep learning with heritage diagnostics can offer a replicable template for other fragile, repetition-rich surface historical materials of architectural heritage.

1. Introduction

1.1 Development History of Shedthin Tiles and Their Application in Ancient Chinese Architecture

Shedthin tiles, also known as "brick sheathing" or "wangba (望芭)", are important components of the roof structure in ancient Chinese architecture. In practical applications, shedthin tiles primarily serve to carry the finish roof tiles, block rainwater infiltration and act as an interior wind- and dust-screen while lending the ceiling a refined, rhythmic appearance (Figs. 1, 2). According to *Yingzao Fayuan* (《营造法原》, The Craft of Architectural Construction), the laying process of shedthin tiles must adhere to strict standards: "When laying shedthin tiles, a face edge (mian-yan, 面檐) should first be set at the eaves, and shedthin tiles should be laid on wooden rafters. A primary bedding (le-wang, 勒望) is set on each purlin to prevent the shedthin tiles from sliding and ensure neat horizontal alignment. For shedthin tiles with insufficient width, narrow infill pieces called "zhao-wang" (找望) should be inserted at the upper end of each bay. If the roof has "flying rafters"(fei-chuan, 飞椽) an additional layer of shedthin tiles should be laid on the flying rafters. To protect shedthin tiles from the heavy pressure of the upper roof, wooden strips are typically laid between the shedthin tiles and the turtle-shell plates to buffer the load" (Hou and Hou, 2014).

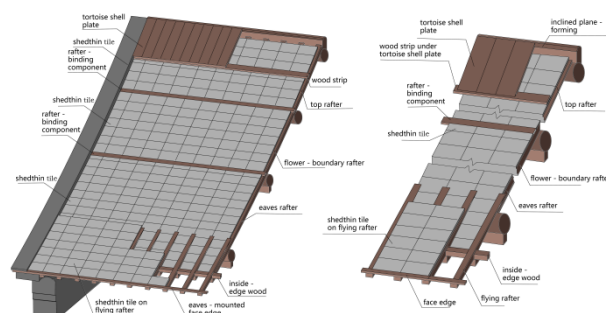


Figure 1. Laying steps and detail drawings of shedthin tiles.

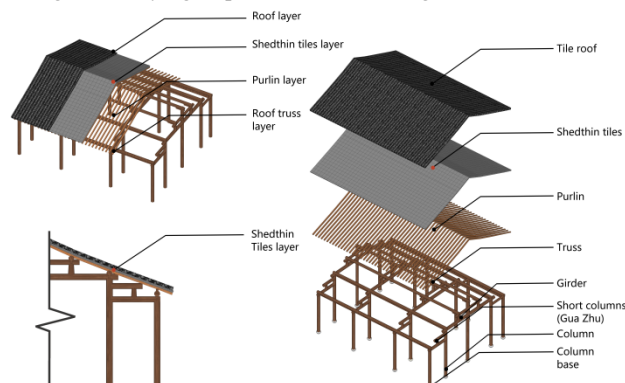


Figure 2. Structural position of shedthin tiles in garden architecture.



Figure 3. Extensive application of shedthin tiles in Suzhou garden architecture.

The application of shedthin tiles in architecture can be traced back to the Song and Jin dynasties. During the Ming and Qing dynasties, their application gradually narrowed, mainly used in small buildings, especially in the Jiangnan region. However, the limitations of ancient manufacturing techniques led to shedthin tiles having the disadvantages of brittle texture and insufficient toughness, making them prone to damage and detachment due to the deformation of wooden rafters, hence they were not widely promoted. In the architecture of Suzhou gardens, shedthin tiles are relatively widely used (Fig. 3). During building repairs, shedthin tiles are often polished and reused, and the differences in their age, color, and material often form a unique texture effect, thereby enhancing the visual beauty of the architecture.

1.2 Material Characteristics and Pathological Types of Shedthin Tiles

Shedthin tiles in the Suzhou area can be divided into three categories according to their craftsmanship: rough shedthin tiles, brushed shedthin tiles, and fine-processed shedthin tiles. Rough shedthin tiles are unprocessed and are mostly used in areas that cannot be directly seen; brushed shedthin tiles are coated with a grayish-white slurry on the surface before laying, and a white edge is trimmed at the corners to form a continuous white line; fine-processed shedthin tiles are finely processed through planing, polishing, etc., and are mostly used in buildings of higher grades.

Compared with wooden sheathing, shedthin tiles have better water absorption performance, thus they can better protect the underlying wooden rafters from decay. Mr. Chai Zejun wrote in The Repair Engineering Report of the Mituo Hall in Chongfu Temple, Shuozhou: "Using shedthin tiles for the hall roof is more conducive to water absorption than wooden sheathing, which is quite advantageous for protecting the rafters from decay. In the northwest corner, a small part of the sheathing was replaced by later generations, and the eaves rafters are severely decayed, which was not discovered in the past" (Chai, 1993). Due to this characteristic of shedthin tiles, they have been widely used in the hot and humid climate of southern China.

The pathologies of shedthin tiles mainly stem from climatic factors and construction techniques. The hot and humid environment in the south and long-term erosion by rainwater keep shedthin tiles in a humid state, making them prone to mold growth and accelerating brick damage. In winter, after multiple freeze-thaw cycles of ice formation from water, the structure of shedthin tiles is easily damaged. In addition, uneven thickness during the firing process and subsequent processing and polishing easily cause structural damage, accelerating their deterioration.

In terms of pathological types, China's current standard Overview of Diseases and Classification of Stone and Brick

Collections (GB/T 30688, 2014) classifies brick and stone pathologies into fifteen categories, including cracks, local defects, surface corrosion, pulverization, lamellar glass, rust crusts, surface pigment damage, etc. In addition, the pathological characteristics of masonry heritage can be broadly grouped into three types: mechanical damage, surface weathering, and surface pollution and discoloration (Zhou et al., 2014). Therefore, considering the above information and combining with the actual situation of Suzhou garden architecture, this paper summarizes the four typical pathologies of shedthin tiles and their characteristics as follows (Table 1):

- (1) Water stains: Under the long-term action of humidity and freeze-thaw cycles, water accumulation in shedthin tiles forms water stains, accelerating aging. Water stains are divided into surface water stains and penetrating water stains, and the presence of water stains often indicates roof leakage problems.
- (2) Color aberration: Inconsistent colors or polishing scratches when treating the surface of shedthin tiles with gray water or mortar often lead to color aberration. Scratched areas easily absorb dust and pollutants, making the brick surface darken or turn yellow.
- (3) Surface scaling: Rainwater and humidity affect the aging of surface mortar, and water penetration forms bubbles. After the bubbles burst, the surface material peels off.
- (4) Excessive gaps: In the repair of ancient buildings, shedthin tiles are often cleaned, polished, and reused after processing (Li et al., 2023). However, due to long-term wear and tear, the size of reused shedthin tiles is reduced, and the brick joints cannot be aligned during laying, affecting waterproof performance and structural stability, and even threatening pedestrian safety.

At present, most of the new gardens and antique buildings use aluminum boards to replace the traditional shedthin tiles, and the number of craftsmen who can fire the shedthin tiles is decreasing, so it is very important to identify and repair the pathologies of the shedthin tiles in time to protect the world cultural heritage of Suzhou's classical gardens.





| water stains | color aberration |
|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
|  |  |
| surface scaling | excessive gaps |
|  |  |

Table 1. Four typical pathology types of shedthin tiles.

1.3 Traditional Pathology Detection workflow and Its Limitations

Traditionally, detecting pathologies in historic buildings relies on manual inspection, following a workflow of survey,

diagnosis, treatment, monitoring, and prognosis. However, this approach has significant limitations. For example, this model relies on on-site visual inspection by experts or data collection through photography and other means, followed by post-analysis to assess the damage status. Although these methods can ensure detection accuracy to a certain extent, they require a lot of human and material resources. Moreover, this method lacks a quantitative assessment of the degree of damage, and the detection results are easily affected by the subjective judgment of inspectors, limiting their rationality. Therefore, there is an urgent need for a more efficient detection method to assist inspectors in accurately identifying damage and reducing unnecessary further damage.

2. Deep-Learning Object Detection for Architectural Heritage Pathologies

2.1 Evolution of Deep Learning-Based Object Detection Technology

With the rapid development of deep learning, deep learning-based object detection technology has been widely used due to its excellent performance. Object detection extracts features from data such as images and videos and uses multi-layer neural networks to achieve learning and recognition of complex patterns. The current mainstream deep learning-based object detection models can be divided into two categories: one-stage object detection models represented by SSD and YOLO, and two-stage object detection models represented by the RCNN series. Compared with two-stage object detection models, one-stage object detection models only need to extract target features once, with the advantage of fast detection speed, but their accuracy is generally slightly lower than that of two-stage models (Zhang et al., 2023).

R-CNN was the first deep learning-based object detection algorithm, which mainly uses a convolutional neural network (CNN) to extract features of target regions and a support vector machine (SVM) for classification (Girshick et al., 2015). The subsequent Fast R-CNN made many improvements based on R-CNN, taking the entire image as input and introducing a region of interest pooling layer (RoI pooling) to extract features (Girshick, 2015). These two-stage detectors divide the detection process into two stages: feature extraction and regression classification, and finally output the results. Although such algorithms can provide higher accuracy, they also bring higher computational requirements (Zhu et al., 2023). Therefore, in order to reduce computational complexity and improve detection efficiency, many one-stage models have emerged. For example, SSD (Single Shot MultiBox Detector) is a one-stage algorithm that can predict bounding boxes on multiple scales to achieve detection of objects of different sizes (Liu et al., 2016). In recent years, in the field of object detection, the YOLO (You Only Look Once) series of algorithms has stood out for its real-time performance and high accuracy. This algorithm transforms the object detection problem into a regression problem and achieves end-to-end object detection through a single convolutional neural network (Wang et al., 2023).

2.2 Applications in Architectural Heritage Pathology Detection

With the development of computer vision technology, especially the development of object detection technology, it has brought new opportunities for the innovation of architectural heritage pathology detection technology (Zhang, 2024). For example, in 2022, Ma Jian, Yan Weidong, et al. used

deep learning models such as YOLO, SSD, and Faster RCNN for crack detection in ancient building timber structures and compared the advantages and disadvantages of different models (Ma et al., 2022). In 2023, Karadag, I. used a conditional generative adversarial network (cGAN) to predict the missing or damaged parts of historical buildings within the scope of early Ottoman tombs (Karadag, 2023).

Since its release in 2016, the YOLO series of algorithms has received widespread attention due to its excellent detection performance and efficient detection speed. In recent years, with the continuous optimization of the YOLO algorithm, the application of the YOLO series of algorithms in architectural heritage conservation has also made great progress, especially in the surface pathology detection of building materials such as bricks and stones, roof tiles, and wooden components (Mishra and Lourenço, 2024). For example, Yan, L. et al. used the YOLOv4 model to detect pathologies of shedthin tiles in Suzhou gardens, achieving an average mAP50 of 43.57% (Yan et al., 2024). Narges Karimi et al. used the YOLOv7 model and trained the model with more than 5,000 collected pathology images to detect pathologies of traditional Portuguese tiles, achieving an overall accuracy of more than 72% (Karimi et al., 2024).

In addition to simply improving traditional pathology detection methods, some scholars have also optimized and improved the object detection model itself to optimize model performance and improve detection efficiency. For example, Qiu et al. used an improved YOLOv8 model to detect roof pathologies in traditional villages in Fujian, with an average accuracy of 89.4%. Compared with the baseline model, this method improved the average accuracy by 1.5% while reducing model parameters and computational complexity (Qiu et al., 2024). In addition, some scholars have also explored unsupervised learning models such as autoencoders for intelligent detection systems to detect roof surface pathologies (Zhang et al., 2024).

In recent years, deep learning-based object detection technology has shown significant advantages in architectural heritage pathology detection. This type of method not only improves detection efficiency but also ensures high accuracy. With the in-depth research and technological development, the detection objects have gradually expanded from single materials to multiple materials, and model performance and detection efficiency have been continuously optimized. In the future, with the continuous update and iteration of technology and the increasing integration of interdisciplinary, deep learning technology will continue to promote innovation and development in the field of architectural heritage conservation.

3. Deep-Learning Workflow for Intelligent Pathology Detection of Shedthin Tiles

3.1 Data Collection

In this study, a total of 1,283 high-resolution pathology photos covering various pathology types were collected from multiple classical gardens in Suzhou (Fig. 4). To ensure data diversity, the shooting locations were selected from different pavilions, terraces, and halls, which are representative and have different pathology conditions. During data acquisition, the team used a mobile-phone camera to photograph the tiles under diverse weather conditions—clear, overcast, and immediately after rainfall—to bolster the model's robustness against environmental and lighting variations.



Figure 4. Data collection locations.

3.2 Data Preprocessing

Through manual screening, a total of 1,250 high-quality photos were retained for dataset construction. Each photograph was then pre-processed—shadow suppression and denoising—to maximise the visibility of pathology features. All images were uniformly rescaled to 512×512 pixels, both to satisfy YOLO's input specification and to reduce computational overhead. Finally, an extensive augmentation pipeline—random cropping, flips, rotations, color jitter, and synthetic noise—was applied to enlarge the effective training set and strengthen the model's generalisation capacity. (Chen et al., 2021).

3.3 Dataset Construction

In the dataset construction process, Labelme was selected because its polygon-based tools allow more accurate delineation of pathology margins—critical for downstream quantitative analysis. In addition, several team members independently produced labels, after which a senior reviewer reconciled discrepancies and ensured stylistic consistency. After the annotation was completed, a computer-vision specialists further verified the annotated data, including geometry, class codes, and file integrity. Finally, this study constructed a high-quality dataset containing 1,250 annotated photos, which was randomly divided into a training set (1,000 photos) and a validation set (250 photos) at an 8:2 ratio to ensure data quality.

3.4 YOLO11-Seg Architecture and Training Configuration

As a representative of one-stage object detection models, YOLO (You Only Look Once) was proposed by Joseph Redmon et al. in 2015, and its core is to achieve end-to-end real-time object detection through regression problem modeling. YOLO11 is the latest object detection model in the YOLO series, further improved and optimized by Ultralytics based on the original YOLOv8 model.

This study adopted the YOLO11-seg model optimized for instance segmentation tasks. The structure of YOLO11-seg mainly includes three parts: the backbone network, the neck network, and the head network (Fig. 5). It uses an improved CSPDarknet in the backbone network, which improves computational efficiency and feature extraction capabilities through cross-stage partial connections (CSP) to ensure that the model can not only process data quickly but also capture rich feature information. In addition, the C3k2 module is introduced in the backbone network to replace the traditional Bottleneck structure, which not only reduces the amount of computation but also maintains efficient feature learning capabilities. In terms of the neck network, the model combines the feature pyramid network (FPN) and the path aggregation network (PAN)

to generate multi-scale feature maps and enhance feature fusion and transmission. The model also adds the C2PSA position-sensitive attention mechanism to further improve the effect (Ali and Zhang, 2024). The head network design of YOLO11-seg mainly includes a decoupled head and a segmentation head optimized for instance segmentation, to support classification, regression, and high-quality pixel-level mask output. In addition, YOLO11-seg uses the SPP module and various attention mechanisms to further improve the model's receptive field and performance.

The YOLO11-Seg model was trained for 362 epochs using a two-phase transfer-learning schedule. This study adopted a transfer learning strategy in the experimental stage. In the first stage, a pre-trained model was loaded, the backbone network was frozen, the learning rate was 0.005, the batch size was 8, and 3 epochs of training were conducted to extract basic features. In the second stage, the learning rate was adjusted to 0.00001, the backbone network was unfrozen, and deep features were extracted. In the final 10 epochs of training, Mosaic augmentation was turned off to improve detection performance in real-world scenarios. During the entire training process, if the model performance did not improve after 100 epochs, the training would automatically stop.

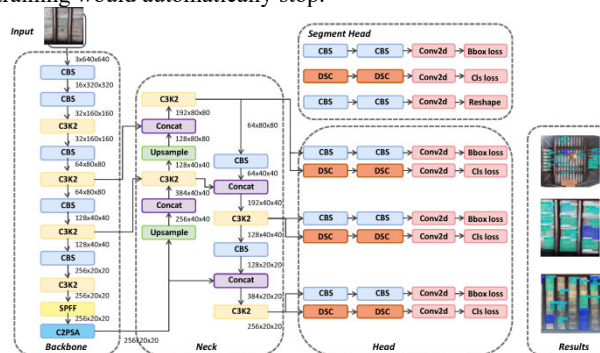


Figure 5. Architecture diagram of YOLO11-seg model.

4. Model Result Testing and Analysis

4.1 Model Testing Results and Manual Verification

After 362 epochs of training, the loss-curve analysis (Fig. 6) revealed that the network reached its minimum object-detection loss at epoch 355, while the instance-segmentation loss bottomed out at epoch 360. Considering both tasks jointly, the aggregate loss was lowest at epoch 360. Accordingly, the weights from epoch 360 were selected for subsequent performance evaluation and quantitative pathology analysis. During testing, the IoU threshold was fixed at 0.5. On the validation set, the model achieved a mean average precision of 51.7 % for object detection (mAP50) and 49.8 % for instance segmentation (Fig. 7). The detection accuracies of different pathologies varied significantly. For example, in the instance segmentation task, the accuracy of excessive gaps (EG) was 69.5%, while that of surface scaling (SS) was only 30.9%. Further analysis of the confusion matrix showed that surface scaling (SS) was easily confused with color aberration (CA) and water stains (WS), which may be due to the diverse and overlapping morphologies of shedthin tile pathologies, thereby increasing the difficulty of model recognition. For example, surface scaling varies in morphology, size, color, and texture, and exists in various forms, some of which are large-area scaling, while others present punctate scaling. This diversity undoubtedly increases the difficulty for the model to accurately recognize pathology features.

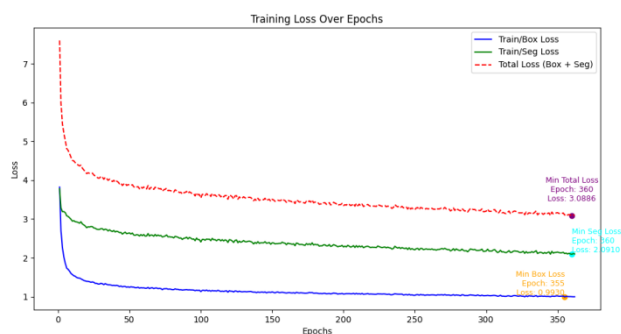


Figure 6. Training loss gradient diagram of YOLO11-seg model.

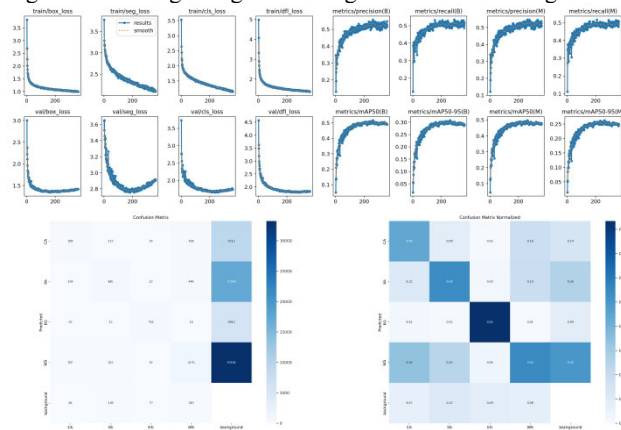


Figure 7. Performance metrics and confusion matrix of YOLO11-seg model.

| Class | Images | Inst- ances | Box | | |
|-------|--------|----------------|---------------|--------|-------|
| | | | Precisi on | Recall | mAP50 |
| all | 236 | 5357 | 0.559 | 0.506 | 0.517 |
| CA | 200 | 898 | 0.558 | 0.529 | 0.558 |
| SS | 190 | 1272 | 0.472 | 0.342 | 0.347 |
| EG | 154 | 851 | 0.69 | 0.669 | 0.694 |
| WS | 234 | 2336 | 0.515 | 0.483 | 0.469 |

| Class | Images | Inst- ances | Mask | | |
|-------|--------|----------------|---------------|--------|-------|
| | | | Precisi on | Recall | mAP50 |
| all | 236 | 5357 | 0.546 | 0.491 | 0.498 |
| CA | 200 | 898 | 0.559 | 0.523 | 0.551 |
| SS | 190 | 1272 | 0.444 | 0.318 | 0.309 |
| EG | 154 | 851 | 0.686 | 0.662 | 0.695 |
| WS | 234 | 2336 | 0.496 | 0.461 | 0.436 |

Table 2. Model detection performance metrics for different pathologies.

Subsequent analysis revealed that shedthin tiles frequently exhibit overlapping deterioration types: for instance, water staining (WS) often coincides with surface scaling (SS). Because YOLO treats every bounding box independently, overlapping annotations cannot be recognised simultaneously; annotators therefore marked only the most visually dominant defect, inevitably depressing detection accuracy. To mitigate this limitation, we introduced the notion of “compound pathologies” (Fig. 8), explicitly labelling overlap zones as combined classes (e.g., WS + SS). This strategy reduces annotation bias and improves model sensitivity to complex, co-occurring damage. Although the current detection performance

still falls short of the ideal, the experiment confirms the method’s promise for non-destructive diagnosis of shedthin-tile pathologies.

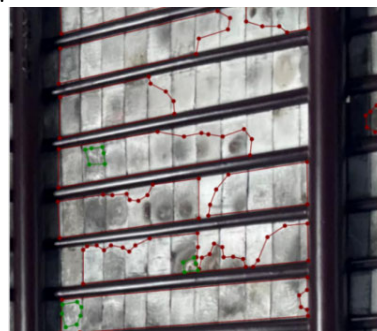


Figure 8. Schematic diagram of compound pathology annotation.

Subsequently, we manually validated the model using a separate set of 566 photographs. The detection standard was that a detection box was considered correctly detected if it overlapped with the annotated box by more than 50% (Table 2). The results showed that the detection accuracies of water stains (WS), color aberration (CA), surface scaling (SS), and excessive gaps (EG) were 79.31%, 73.38%, 61.12%, and 75.60%, respectively, with an average accuracy of 74.38% (Fig. 9). Among them, the detection performance of surface scaling (SS) was relatively weak, providing a clear direction for model optimization. Subsequent model improvements will enhance its robustness through data augmentation, attention mechanism improvements, or multi-scale feature fusion.

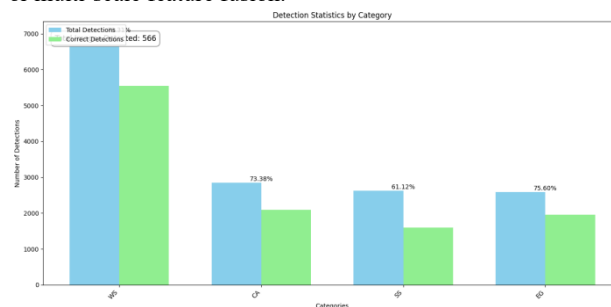


Figure 9. Model detection accuracy chart.

4.2 Model Application and Analysis

In the application phase, complete roof-elevation panoramas were extracted from two Suzhou gardens—Mountain Villa with Embraced Beauty (环秀山庄) and the Garden of the Couple’s Retreat (耦园)—and processed through the trained pipeline (Fig. 10). The analysis proceeded as follows:(a)Site localisation and roof extraction The target buildings were pinpointed within each garden; their roof typology and the position of the shedthin-tile layer were confirmed, and a seamless elevation image was generated;(b)Pathology detection The YOLO11-Seg model was applied to the elevation to produce pixel-level masks of each deterioration type;(c)Thermal mapping Segmentation outputs were aggregated into heat maps, revealing the spatial distribution of roof pathologies;(d)In the application phase, complete roof-elevation panoramas were extracted from two Suzhou gardens—Mountain Villa with Embraced Beauty and the Garden of the Couple’s Retreat—and processed through the trained pipeline (Fig. 10). The analysis proceeded as follows:Quantitative ranking Coverage and lesion-area statistics were computed for every class, allowing the relative severity of shedthin-tile pathology across the roofscape to be objectively ranked.

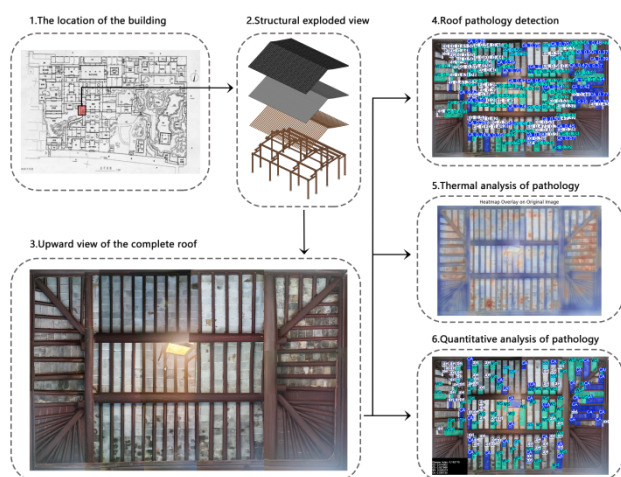


Figure 10. Basic flowchart of model application analysis.

The pathology detection results showed (Fig. 11) that the model used in this study performed ideally in detecting shedthin tile pathologies of complete roofs, basically achieving the research expected goals. At the same time, the pathology detection heatmap displayed the response intensity distribution of the model in different roof areas, where the red area represented the area with high model response values, and relatively, the blue area indicated the background area with low response or non-key areas. By observing the heatmap, it can be found that although the overall model performance is satisfactory, there are still certain missed detections and false detections in some cases, and the accuracy still has room for improvement.

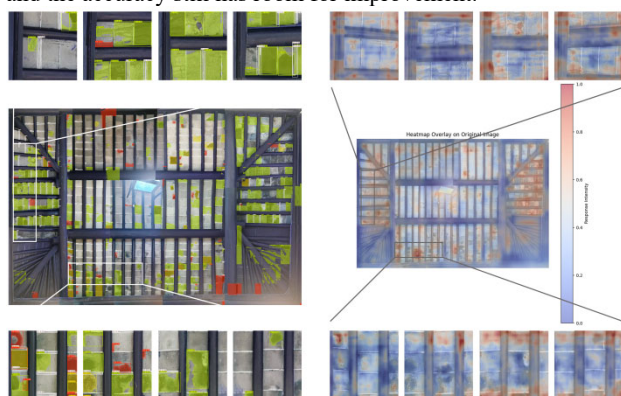


Figure 11. Pathology detection results and thermal analysis chart.

In terms of quantitative analysis, the Guidelines for Safety Assessment of Modern Historical Building Structures (WW/T 0048, 2014) propose a structural damage assessment method based on the proportion of damaged area, which provides a reference for the quantification of architectural heritage pathology severity. In addition, some scholars have quantified the damage severity of heritage buildings or sites by calculating the proportion of pathology area or conducting weighted assessments of pathology severity (Zhai, 2024; Yao and Sun, 2016). To express the overall deterioration state of the shedthin-tile layer on a given roof, we define a dimensionless indicator—the Shedthin-Tile Pathology Index (STPI). The calculation formula for this index is:

$$D = \frac{S_{ca}}{S_t} \alpha + \frac{S_{ws}}{S_t} \beta + \frac{S_{ss}}{S_t} \gamma + \frac{S_{eg}}{S_t} \delta$$

In this formula, S_{xx} represents the area of each type of pathology, S_t is the total area of the roof, and α , β , γ , δ are dimensionless weight coefficients, whose specific values are determined

according to the severity of different pathologies on component hazards. In this study, according to the hazard levels of these four pathologies, the coefficients were determined as $\alpha=0.5$, $\beta=1$, $\gamma=2$, $\delta=1.5$, respectively. Based on this formula, the shedthin tile pathology index of a building can be calculated. By comparing the shedthin tile pathology indices of different buildings, the relative severity of shedthin tile component damage can be determined to set the priority for subsequent repairs.

Quantitative assessment was carried out on two garden-roof elevations. The resulting Shedthin-Tile Pathology Index values were 0.162175 for Roof 1 and 0.093364 for Roof 2 (Fig. 12,13). Because Roof 1 exceeds the 0.15 threshold, it ranks as the higher-priority candidate for intervention, whereas Roof 2, with a markedly lower index, can be scheduled for subsequent treatment. Incorporating these index scores into the repair timetable supports rational resource allocation and ensures that conservation funds are directed first to the areas of greatest need.

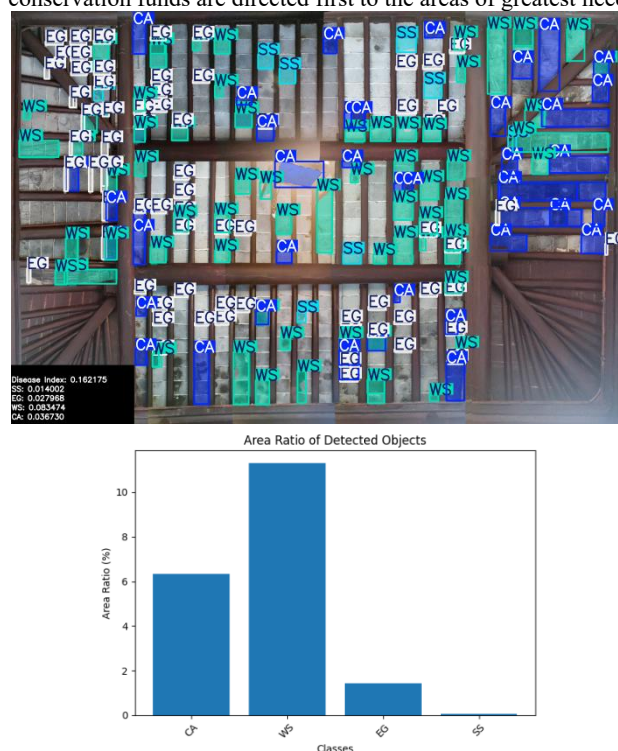


Figure 12. Quantitative detection results and pathology area ratio of roof 1.



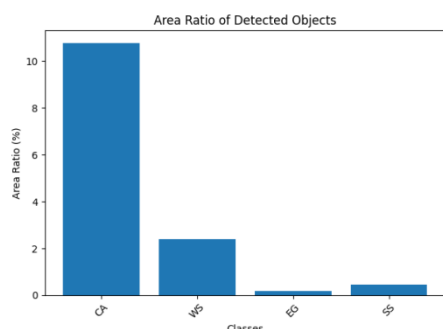


Figure 13. Quantitative detection results and pathology area ratio of roof 2.

4.3 Inter-model Performance Comparison

Before contrasting the various YOLO versions, an internal benchmark was essential for three reasons: it eliminates confounding variables by training and testing every model under identical data, hardware, and hyper-parameter conditions; it verifies whether architectural upgrades from YOLOv8 through YOLO11 actually translate into measurable gains for the specific task of shedthin-tile pathology detection; and, most pragmatically, it pinpoints the model that offers the best balance of accuracy, inference speed, and resource consumption for eventual field deployment.

A head-to-head evaluation was conducted on four segmentation-enabled YOLO variants: YOLOv8-seg, YOLOv9-seg, YOLOv10-seg, and YOLO11-seg. Validation-set results (Fig. 14) reveal that, aside from YOLOv9-seg model, which had significantly lower detection accuracy due to technical problems during training, the other four models showed relatively close detection accuracies, with the difference in Box-mAP50 and Mask-mAP50 indicators within 3%, indicating that each model has high performance consistency in object detection and instance segmentation tasks.

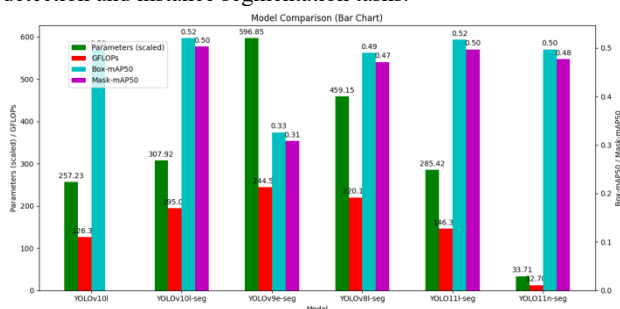


Figure 14. Comparison of detection results of different models.

Among them, the YOLO11n-seg model stands out particularly in terms of parameter quantity and computational efficiency. Experimental data showed that the number of parameters and floating-point operations (GFLOPs) of this model were both less than 1/10 of other models, significantly lower than other models. This characteristic enables YOLO11n-seg to ensure faster inference speed, lower hardware requirements, and make subsequent integration with real-time terminal devices such as drones possible, thereby significantly expanding its application potential.

5. Conclusions and Discussions

Shedthin tiles constitute both a functional weather-shield and a signature aesthetic element within Suzhou classical gardens.

Their conservation is increasingly threatened by environmental exposure, material ageing, and the dwindling transmission of traditional craftsmanship. To address the low efficiency and limited quantitative insight of manual surveys, this study developed a deep-learning pathology-detection workflow. The study uses the YOLO11-seg model, and by constructing a high-quality dataset containing 1,250 high-resolution sample images, based on the model training completed with 362 iteration cycles, the final mean average precision (mAP50) of 51.7% is achieved. Through manual verification, the model's automatic recognition accuracies for the four main pathologies of water stains, color aberration, scaling, and excessive gaps reached 79.31%, 73.38%, 61.12%, and 75.60%, respectively, with a comprehensive accuracy of 74.38%, showing great application potential. The experimental results show that this method has the following significant advantages:

- (1) Improving the efficiency of building pathology detection: Compared with traditional manual detection, deep learning technology significantly accelerates the detection speed of architectural heritage pathologies, plays an important role in large-scale detection, and thus saves a lot of human and material resources.
- (2) Enhancing the detection accuracy of shedthin tile components: This study uses the advanced YOLO11-seg model to further improve the detection accuracy of shedthin tile components based on previous research.
- (3) Realizing quantitative analysis of shedthin tile pathologies: Using the YOLO11-seg model, this study can effectively segment detected targets, realizing quantitative analysis of pathology detection. At the same time, the study introduces the index of "shedthin tile pathology index", providing an important basis for determining the priority of subsequent repair work and laying the foundation for the practical application of the technology.

While this study highlights the promise of deep learning for pathology detection in architectural heritage, notable limitations remain in dataset coverage, model precision, and operational robustness. We therefore identify four parallel lines of future work:

- (1) Constructing a higher-quality dataset: Future research will collect more pathology data and consider different data sources and environmental conditions to further enhance the model's generalization ability and adaptability.
- (2) Improving the model to enhance detection accuracy: Researchers will continue to optimize and improve the model in follow-up studies, introducing different mechanisms to further enhance detection accuracy.
- (3) Considering multi-modal learning: Combining deep learning with multi-modal detection technologies such as point cloud technology and laser scanning to improve the model's detection performance and applicability.
- (4) Deploying real-time detection terminals: Exploring the combination with technologies such as drone photography, deploying real-time detection terminals, and realizing real-time monitoring of building pathologies.

Future research will focus on resolving the current technical constraints and practical limitations, and promoting the application of deep learning technology in non-destructive detection of architectural heritage pathologies through continuous model optimization and integration with other technologies. Against the backdrop of the rapid development of artificial intelligence technology, introducing deep learning technology into the field of heritage conservation will further promote interdisciplinary integration and promote the traditional

heritage conservation field to shift from experience-based to technological and innovative.

References

- Ali, M. L., & Zhang, Z. 2024. "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection." *Computers* 13 (12): 336.
- Chai, Zejun. 1993. *Repair Engineering Report of the Mituo Hall in Chongfu Temple, Shuozhou*. Cultural Relics Press. Pp. 64–67.
- Chen, H., Chen, J., & Ding, J. 2021. "Data Evaluation and Enhancement for Quality Improvement of Machine Learning." *IEEE Transactions on Reliability* 70 (2): 831–847.
- GB/T 30688 – 2014. 2014. *Overview of Diseases and Classification of Stone and Brick Collections*. Standard Press of China.
- Girshick, R. 2015. "Fast R-CNN." In *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. 2015. "Region-based Convolutional Networks for Accurate Object Detection and Segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38 (1): 142–158.
- Hou, Hongde, and Hou Xiaoyi. 2014. *Illustrated Treatise on the Construction Principles of "Yingzao Fa Yuan"*. China Architecture & Building Press. Pp. 224–225.
- Karadag, I. 2023. "Machine Learning for Conservation of Architectural Heritage." *Open House International* 48 (1): 23–37. doi: 10.1108/OHI-05-2022-0124.
- Karimi, N., Mishra, M., & Lourenço, P. B. 2024. "Deep Learning–Based Automated Tile Defect Detection System for Portuguese Cultural Heritage Buildings." *Journal of Cultural Heritage* 68: 86–98. DOI: 10.1016/j.culher.2024.05.009.
- Li, Hengyu, Shen Weijun, and Zhu Lei. 2023. "The Renovation Technology of Sheathing Brick of Ancestral Temples at Huishan Creek in Wuxi." *Ancient Building and Garden Technology* (2): 78–82.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. 2016. "SSD: Single Shot Multibox Detector." In *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 – 14, 2016, Proceedings, Part I*, 21–37. Springer.
- Ma, Jian, Yan, Weidong, and Liu, Guoqi. 2022. "Complex Texture Contour Feature Extraction Method of Cracks in Timber Structures of Ancient Architecture Based on Deep Learning." *Journal of Shenyang Jianzhu University (Natural Science)* 38 (5): 896–903.
- Mishra, M., & Lourenço, P. B. 2024. "Artificial Intelligence–Assisted Visual Inspection for Cultural Heritage: State-of-the-Art Review." *Journal of Cultural Heritage* 66: 536–550.
- Qiu, H., Zhang, J., Zhuo, L., et al. 2024. "Research on Intelligent Monitoring Technology for Roof Damage of Traditional Chinese Residential Buildings Based on Improved YOLOv8: Taking Ancient Villages in Southern Fujian as an Example." *Heritage Science* 12: 231. <https://doi.org/10.1186/s40494-024-01345-8>.
- Wang, Linyi, Bai, Jing, and Li, Wenjing, et al. 2023. "Research Progress of YOLO Series Target Detection Algorithms." *Computer Engineering and Applications* 59 (14): 15–29.
- WW/T 0048–2014. 2014. *Guidelines for Safety Assessment of Modern Historical Building Structures*.
- Yan, L., Chen, Y., Zheng, L., et al. 2024. "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." *Heritage Science* 12: 72. <https://doi.org/10.1186/s40494-024-01185-6>.
- Yao, Xue, & Sun, Manli. 2016. "The Quantitative Evaluation of Deterioration Degrees of Earthen Sites Based on Gray Correlation Analysis." *Dunhuang Research* (1): 128–134. DOI: 10.13584/j.cnki.issn1000-4106.2016.01.017.
- Zhai, Qingyi. 2024. "Research on the Detection Method of Moss Growth Rate on Stone Cultural Relics Based on Swin Transformer." Dissertation, Chongqing University of Technology. DOI: 10.27753/d.cnki.gcqgx.2024.000135.
- Zhang, Jiahao. 2024. "Research on Cultural Heritage Disease Monitoring Based on Artificial Intelligence and Experimental Measurement." *World Architecture* (11): 98 – 99. DOI: 10.16414/j.wa.2024.11.019.
- Zhang, Y., Kong, L., Fordjour, M., & Zhang, Q. 2024. "An Integrated Method Using a Convolutional Autoencoder, Thresholding Techniques, and a Residual Network for Anomaly Detection on Heritage Roof Surfaces." *Buildings* 14 (9): 2828. <https://doi.org/10.3390/buildings14092828>.
- Zhang, Yangting, Huang Deqi, Wang Dongwei, et al. 2023. "Review on Research and Application of Deep Learning-Based Target Detection Algorithms." *Computer Engineering and Applications* 59 (18): 1–13.
- Zhou, Weiqiang, Zhou Ping, and Wang Yongjin. 2014. "Overview of Pathologies and Classification of Brick and Stone Cultural Relics." *Cultural Relics* (6): 73–75.
- Zhu, Hao, Zhou, Shunying, and Liu, Xue, et al. 2023. "Survey of Single-Stage Object Detection Algorithms Based on Deep Learning." *Industrial Control Computer* 36 (4): 101–103.