

RESEARCH

Open Access



# Advancing architectural heritage: precision decoding of East Asian timber structures from Tang dynasty to traditional Japan

Chan Gao<sup>1</sup>, Genfeng Zhao<sup>2</sup>, Sen Gao<sup>3</sup>, Shuxin Du<sup>4,5</sup>, Eunyoung Kim<sup>1</sup> and Tao Shen<sup>6\*</sup>

## Abstract

The convergence of cultural and aesthetic elements in timber structures from China's Tang Dynasty (618–907 AD) and traditional Japanese architecture provides a rich tapestry of architectural evolution and cross-cultural exchanges. Addressing the challenge of distinguishing and understanding the intricate styles of these structures is significant for both historical comprehension and preservation efforts. This research introduces an innovative approach by integrating the Multi-Head Attention (MHA) mechanism into the YOLOv8 model, enhancing the detection of architectural features with improved precision and recall. Our novel YOLOv8-MHA model not only demonstrates a notable improvement in recognizing intricate architectural details but also significantly advances the state of the art in object detection within complex settings. Quantitative results underscore the model's effectiveness, achieving a precision of 95.6%, a recall of 85.6%, and a mean Average Precision (mAP@50) of 94% across various Intersection over Union (IoU) thresholds. These metrics highlight the model's superior capability to accurately identify and classify architectural elements, especially within environments rich with nuanced details, utilizing the enhanced YOLOv8-MHA algorithm. The application of our model extends beyond mere architectural analysis; it offers new insights into the intricate interplay of cultural identity and adaptability inherent in East Asian architectural heritage. The study establishes a solid foundation for the meticulous classification and analysis of architectural styles in timber structures within an expansive cultural and historical context, thereby enriching our understanding and preservation of these traditions.

**Keywords** Cultural heritage conservation, Historic timber structures, Tang dynasty architecture, Japanese traditional architecture, Object detection, YOLOv8 algorithm

## Introduction

### Overview of timber structures of Tang dynasty architecture in China and traditional timber structures in Japan

The architectural journey from the Tang Dynasty in China (618–907 AD) to Japanese traditional architecture showcases a curtain of cultural exchange and aesthetic assimilation, with timber structures playing a pivotal role. The Tang Dynasty in China, a pinnacle of cultural and artistic achievement, was renowned for its timber structures that featured a grand scale, symmetrical layout, and intricate decorations [1]. These timber structures transcended regional boundaries and profoundly influenced neighboring countries, especially Japan. Driven by diplomatic and cultural exchanges, such as the dispatch of

\*Correspondence:

Tao Shen  
shentao@tongji.edu.cn

<sup>1</sup> Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology, Nomi 9231292, Japan

<sup>2</sup> Hangzhou R&F Real Estate Development Co., Ltd, Hangzhou 310000, Zhe Jiang Province, China

<sup>3</sup> Baoding Shengyang Real Estate Development Co., Ltd, Baoding 071000, Hebei Province, China

<sup>4</sup> Huzhou University, Huzhou 313000, Zhe Jiang Province, China

<sup>5</sup> Huzhou Key Laboratory of Intelligent Sensing and Optimal Control for Industrial Systems, School of Engineering, Huzhou University, Huzhou 313000, China

<sup>6</sup> College of Design and Innovation, Tongji University, Shanghai 200092, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

missions to the Tang Dynasty, elements of Tang Dynasty timber structures were integrated into Japanese timber structures, highlighting a high degree of stylistic convergence rooted in a common cultural paradigm [2].

However, subtle differences exist in the architectural timber structures of each country. Chinese Tang Dynasty timber structures are characterized by complex roof structures such as heavy eaves and flying eaves, coupled with decorations featuring uniquely Chinese motifs like dragons and phoenixes [3]. In contrast, Japanese traditional timber structures, while echoing the structural spirit of Chinese Tang architecture, incorporate local cultural elements, reflected in more understated roof designs and minimalist decorations [4].

The choice of materials and construction techniques for timber structures further accentuates their uniqueness. The abundance of timber resources and exquisite carpentry skills in China were hallmarks of Tang-style timber structures, while Japanese traditional timber structures reflected local environmental adaptations, especially in terms of earthquake resistance.

Fundamentally, while Tang Dynasty timber structures and Japanese traditional timber structures share profound similarities, their differences in details, materiality, and spatial organization eloquently narrate a story of cultural identity and adaptability, enriching our understanding of East Asian architectural heritage and cross-cultural exchanges [5–7].

### The importance and challenges of differentiating these architectural timber structures styles

Comparative studies of Tang Dynasty timber structures and Japanese traditional timber structures reveal the historical context of Sino-Japanese cultural exchange, offering insights into the evolution and adaptation of transnational architectural styles. This endeavor is crucial for understanding the narrative of East Asian architecture but is also fraught with challenges:

- As a portal of cultural exchange: This comparative analysis goes beyond architectural aesthetics, providing a perspective to observe broader cultural, religious, and artistic exchanges between the Tang Dynasty and Japan.
- Evolution and adaptation of timber structures architectural styles: The study reveals the journey of architectural styles, unveiling the subtle nuances of cultural adaptation and stylistic evolution as Tang Dynasty timber structures were integrated into the Japanese timber structures landscape.
- Insights into techniques and materials: It also offers a platform for an in-depth study of ancient architectural techniques for timber structures, reflecting how

geographical and climatic differences influence the choice of building materials and structural design.

The challenges of this academic pursuit include the scarcity and preservation of historical timber structures, understanding the cultural nuances embedded in architectural elements, and the interdisciplinary nature of the research, requiring a fusion of knowledge in architecture, history, art history, and cultural studies.

### Introduction to YOLOv8 and its relevance in architectural analysis of timber structures

YOLOv8 [8–12], the latest version in the YOLO (You Only Look Once) series, is a state-of-the-art deep learning model that has revolutionized real-time object detection tasks. Renowned for its exceptional processing speed and accuracy, YOLOv8 is particularly suited for identifying complex architectural features of timber structures, aiding in the differentiation between Tang Dynasty timber structures and Japanese traditional timber structures. Its capabilities for automatic feature recognition, efficient processing of large-scale image data of timber structures, and serving as a nexus for interdisciplinary research make YOLOv8 a key tool in architectural analysis, paving new pathways for understanding architectural nuances of timber structures through the lens of advanced computational technology.

### Novelty

This research introduces a significant innovation in the application of deep learning to the field of architectural heritage conservation by integrating the Multi-Head Attention (MHA) mechanism into the YOLOv8 architecture. This integration represents a novel approach to enhancing object detection models specifically tailored to recognize complex architectural features, which is critical for distinguishing between the intricate timber structures of the Tang Dynasty and traditional Japanese architecture.

#### Quantitative improvements:

- Precision and Recall: The enhanced YOLOv8-MHA model demonstrates a precision increase from 89.7% to 95.6% and a recall improvement from 84.3% to 85.6% when compared to the standard YOLOv8 model in detecting architectural features within our dataset. This indicates a more accurate identification and reduction in false positives, essential for detailed architectural analysis.
- Mean Average Precision (mAP@50): The model achieves an mAP@50 of 94%, a significant rise from the 88% achieved by the baseline YOLOv8 model. This improvement underscores the model's

enhanced ability to maintain precision across varied detection thresholds, providing more reliable results in practical applications.

- **Detection in Complex Environments:** With the addition of MHA, the YOLOv8-MHA model shows superior performance in scenarios characterized by complex multi-layered backdrops and foreground occlusions, which are common in images of dense urban heritage environments. The model's ability to maintain high accuracy in such conditions is critical for its application in cultural heritage conservation, where such settings are prevalent.

## Research objectives and paper structure

### Research objectives

This study aims to enhance the performance of Tang Dynasty timber structures and Japanese traditional timber structures architectural target detection by integrating the Multi-Head Attention (MHA) mechanism into the YOLOv8 model. Specific objectives include:

- **Enhancing feature recognition accuracy for timber structures:** Introducing the MHA mechanism to bolster the model's ability to recognize architectural details of timber structures, thereby improving classification accuracy.
- **Optimizing model performance for timber structures analysis:** Leveraging the advantages of the MHA mechanism to enhance the model's efficiency and accuracy in processing complex image data of timber structures.
- **Advancing interdisciplinary research on timber structures:** Combining computer vision technology with multidisciplinary knowledge in architecture, history, and art studies to inject new perspectives and methods into the field of ancient architectural research on timber structures.

In this study, the YOLOv8-MHA model excelled in recognizing Tang Dynasty timber structures and Japanese traditional timber structures. The model achieved a precision of 95.6%, a recall rate of 85.6%, an mAP@50 of 94%, and an mAP@50–95 of 80.4%, demonstrating its robust ability to maintain high detection accuracy across different IoU thresholds for timber structures. These results highlight the potential of the YOLOv8-MHA model in precisely identifying and classifying architectural features, especially in environments that require detailed recognition and the integration of multifaceted features in timber structures.

### The paper structure is designed as follows

- **The introduction** includes presenting the research background, outlining the characteristics of Tang Dynasty timber structures and Japanese traditional timber structures architecture, the importance of differentiating these architectural timber structures styles, and an introduction to YOLOv8 and its relevance in architectural analysis.
- **The literature review** discusses the differences between Tang Dynasty and Japanese traditional timber structures, previous methods used for architectural style analysis and their limitations, and the application and research gaps of the YOLO algorithm in architectural style classification of timber structures.
- **Methodology:** Detailed description of data collection, preprocessing, model training, and optimization processes, with particular emphasis on the integration of the MHA mechanism and its role in enhancing the model's performance for timber structures.
- **Results and Discussion:** Presentation of research findings, analysis of model recognition and classification accuracy for timber structures, and discussion of the application of the YOLOv8-MHA model in architectural style recognition.
- **Conclusion and Future Work:** Summarizing research findings, discussing the limitations of the model and research, and proposing suggestions for future research directions in the study of timber structures within the context of Tang Dynasty and Japanese traditional architecture.

This comprehensive approach, integrating advanced computational analysis with a deep understanding of historical and cultural contexts, offers a new perspective on the study of timber structures in East Asian architecture, emphasizing the significance of preserving and appreciating these architectural treasures.

## Literature review

### Differences between timber structures of Tang dynasty architecture and Japanese traditional architecture

Firstly, a defining characteristic of Chinese Tang Dynasty timber structures is their symmetry and central axial layout, prominently seen in imperial palaces, temples, and official buildings. This architectural approach emphasizes a sense of ceremony and authority. The roof designs of Tang Dynasty timber structures typically feature heavy eaves with graceful curves, often adorned with mythical animals like lions and phoenixes on the ridges, symbolizing power and sanctity. Additionally, Tang Dynasty

timber structures are known for their colorful paintings, carvings, and ceramic tiles, showcasing the era's exquisite craftsmanship and attention to detail [13–17].

In contrast, while Japanese traditional timber structures have absorbed many elements from Chinese Tang Dynasty architecture, they also exhibit distinctive features due to geographical and cultural differences. Japanese timber structures retain the central axial symmetry and heavy eaves design, but with more simplified details. For instance, the curvature of the roofs is usually not as pronounced as in China, and the decorations are subtler and more modest. The spatial layout in Japanese timber structures, especially in temples, tends to be more compact, harmonizing with the tranquil natural surroundings. Additionally, traditional Japanese architectural elements such as tatami mats, sliding doors (shoji), and wooden verandas (engawa) distinguish it from Chinese Tang Dynasty architecture, reflecting the Japanese pursuit of functionality and minimalist aesthetics [4, 18–21].

Overall, both Chinese Tang Dynasty timber structures and Japanese traditional timber structures are remarkable representations of their respective cultures. Their similarities and differences narrate a rich story of cultural exchange that transcends time and space [2, 5, 22–27]. The specific distinctions between the two highlight the unique architectural innovations and cultural significance of each tradition.



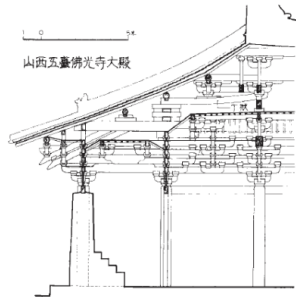



- (1) Tang dynasty architecture often features stone-built platforms (also known as “tai ming”) at the base, while Japanese styles utilize suspended, high wooden platforms. (Fig. 1a)
- (2) The terminations of columns in Tang Dynasty architecture are fashioned into gentle curves or broken lines, imparting a full and soft appearance, a feature referred to as “Juan Sha.” In Japanese traditional architecture, the curvature or broken line design of the columns is more pronounced, and this treatment is often applied to both the top and bottom of the columns.
- (3) The primary color scheme in Tang dynasty style is red and white, occasionally with green doors and windows, while Japanese architecture predominantly uses brown (chocolate color).
- (4) The “dougong” (bracket sets) in Tang-style often feature split-bamboo or qin-face style, whereas Japanese “dougong” typically have uniform ends. (Fig. 1b).
- (5) Tang-style uses flat, straight rafter covers, while Japanese style employs curved rafter covers.
- (6) The wing corner rafters in Tang-style are fan-shaped and laid flat, whereas in Japanese style, they are straight and laid flat.

- (7) Tang-style roofs often use ceramic tiles, while Japanese temples typically use bark (such as cypress bark or thatch) for roofing. (Fig. 1c).
- (8) For facade decoration, Tang-style commonly uses suspended fish ornaments (“xuan yu”), while Japanese style uses raked grass (“re cao”) and often includes metal wind boards (“bo feng ban”). (Fig. 1d).
- (9) Ridge decorations in Tang-style are typically sharp-pointed and made of ceramic (“chi wei”), while Japanese style often uses boot-shaped metal ridge ornaments (gilded or made of copper). (Fig. 1e).
- (10) Tang-style frequently uses straight-lattice windows, while Japanese style predominantly features grid windows.
- (11) The main wooden frame in Tang-style buildings is often of the lifted-beam type, while in Japanese style, it is typically the penetrating tie-beam type “small house group.”
- (12) The roof surface in Tang-style elegantly curves (“ju zhe”), while in Japanese style, the small house group's roof slope is rigid without lifting. (Fig. 1f)
- (13) The roof surface in Tang-style gently curves, following a natural trajectory, while in Japanese style, the beam frame is elevated, resulting in steep roof slopes.

#### Previous methods for wooden structures analysis and their limitations

Research on the application of YOLO models in detecting various aspects of wooden structures has shown promising results. YOLO technology, known for its speed and accuracy, has been utilized in detecting cracks in timber structures of ancient architecture, surface knots in wood, and even wood pith location in cross-sectional images [28–30]. These studies have highlighted the effectiveness of different YOLO versions such as YOLO v5s, YOLO v3, and Tiny-YOLO in accurately identifying defects and features in wooden materials, showcasing their potential in intelligent maintenance, grading, and quality assessment of timber structures. Additionally, the use of YOLO models in detecting building components like doors and windows for energy performance analysis further demonstrates the versatility and applicability of YOLO algorithms in various domains, including the recognition of building wooden structures [31].

In papers on architectural style analysis, Zhe Cui's “Automatic Classification of Ancient Architectural Components Based on Deep Learning” adopted a deep learning-based method to classify point clouds of ancient buildings and a method based on statistical

<p>a</p>		
	<p>Tang Dynasty Architecture with Stone platform("tai ming")</p>	<p>Suspended high bed wooden platform</p>
<p>b</p>	 <p>Source: Drawing of the Main Hall of Foguang Temple, Mount Wutai, Shanxi. From Liang Sicheng's Complete Works, Volume 7 - "Annotations on Architectural Methods," Large Woodwork, 91b, page 140.</p> <p>Precise structural interlocking relationships, load-bearing structures, and massive dougong. In Tang dynasty architecture, the dougong ("bracket sets") often feature split-bamboo style or qin-face style.</p>	 <p>Kiyomizu-dera features delicate dougong ("bracket sets"), primarily serving a decorative purpose. Beyond the dougong, which are sometimes purely ornamental, components such as pseudo rafters are common. The Japanese style often employs uniform-end dougong.</p>
<p>c</p>	 <p>Tang Dynasty roofs often feature ceramic tiles made of pottery.</p>	 <p>Japanese architecture often uses bark (such as cypress bark or thatch) for roofing.</p>

**Fig. 1** Differences between Tang Dynasty Architecture and Japanese traditional Architecture






<p>d</p>	 <p>Tang Dynasty facade decorations often feature suspended fish ornaments.</p>	 <p>Japanese architecture uses grass ("re cao") and metal wind boards with inlaid designs("bo feng ban").</p>
<p>e</p>	 <p>Tang Dynasty ridge decorations often feature sharp-pointed ceramic ridge ornaments("chi wei").</p>	 <p>In Japan, boot-shaped metal ridge ornaments (gilded or made of copper) are commonly used.</p>
<p>f</p>	 <p>Tang Dynasty roofs feature elegantly curved lifting frames ("ju zhe").</p>	 <p>In Japanese style, the roof slopes of the small house group are stiff and straight, without lifting frames.</p>
	<p>Tang Dynasty Architecture</p>	<p>Japanese traditional Architecture</p>

Fig. 1 continued

information for more refined classification [32]. Qiang Ling's research "Classification Study of Knowledge on the Protection of Ancient Architectural Cultural Heritage" utilized text classification techniques, including Bayesian, k-nearest neighbors, and Support Vector Machine (SVM) algorithms [33]. Zengli Shi's "Types and Classification of Ancient Chinese Architecture" systematically categorized ancient Chinese architecture through a structured approach and hierarchical framework system [34]. This method thoroughly reviewed the descriptive features and system structure of architectural types. Jose Llamas and others mainly applied deep learning technology to classify images of architectural cultural heritage, with the key technology being Convolutional Neural Networks (CNNs), used to extract useful information from images [35]. These studies have achieved notable success in their respective fields but also have certain limitations.

Although the above studies provide strong insights into architectural style analysis, these methods often do not specifically analyze the two stylistically similar categories of Tang Dynasty architecture and Japanese traditional architecture. These two categories of architecture have high similarity and complexity in structure, decoration, and cultural significance, presenting a huge challenge for classification algorithms. Existing methods often struggle to handle these subtle yet significant differences, limiting their application in a broader cultural and historical context.

Furthermore, most existing studies have not utilized YOLOv8, the latest one-stage object detection algorithm. Compared to previous algorithms, YOLOv8 shows significant improvements in real-time performance, detection accuracy, and model generalization ability. This is particularly important for architectural style analysis, as the algorithm needs to precisely identify and classify subtle architectural elements and style features. The high real-time performance of YOLOv8 ensures rapid processing of a large volume of image data, while its high-precision detection capability ensures accurate capture of architectural style details, which is especially crucial for distinguishing between such similar styles as Tang Dynasty and Japanese traditional architecture.

In summary, this study, by adopting the YOLOv8 algorithm, not only enhances the understanding of the subtle differences between Tang Dynasty timber structures architecture and Japanese traditional timber structures architecture but also pushes forward the technology of architectural style analysis, laying the groundwork for future precise classification and analysis of architectural styles in a broader cultural and historical context.

## Methodology

### Data collection and data preprocessing

#### Data collection

Famous Tang Dynasty buildings in China (Fig. 2): Examples of Chinese Tang Dynasty architecture include the East Main Hall of Foguang Temple, Daming Palace, Famen Temple, Dacien Temple, Linde Hall (ruins), Tang City Ruins, South Chan Temple Main Hall of Mount Wutai, Tiantai An Main Hall of Pingshun, and the Main Hall of Guangren Temple in Ruicheng.

Japanese traditional buildings (Fig. 3) include: Todaiji Temple, Horyuji Temple, Heijo Palace (site), Tangshodaiji Temple, Kofukuji Temple, Yuanxingji Temple, Ninwaji Temple, Nijo Castle, Kiyomizu Temple, Kinkakuji Temple, Ginkakuji Temple, etc.







#### Data preprocessing and augmentation strategies

To address the lack of datasets for Tang Dynasty architecture in China and Tang-style architecture in Japan, we compiled a unique dataset containing 850 images sourced from the internet and our own photography. To enhance the dataset and prepare for effective model training, we employed data augmentation techniques. These techniques not only expanded the size of the dataset but also introduced variability and complexity, simulating real-world conditions, thereby enhancing the model's robustness and generalizability. The augmentation strategies used are as follows:

*Random augmentation (Fig. 4a)* This technique randomly alters images, such as rotation, flipping, and color adjustment. The goal is to mimic a variety of possible visual scenes that the model might encounter, thereby enriching the dataset and enhancing the model's generalization ability for different visual inputs.




*Mixup (Fig. 4b)* Mixup creates synthetic images by blending a pair of images and their corresponding labels. This method not only expands the dataset but also introduces a regularization effect in the model. It encourages the model to be less confident about the interpolated feature vectors and labels, promoting better generalization and smoother decision boundaries between classes.

*Mosaic (Fig. 4c)* Mosaic augmentation merges four different images into one composite image. This technique is particularly beneficial for object detection tasks, as it increases the frequency of small object appearances and introduces complex backgrounds. The generated composite images help the model better detect objects in dense and cluttered scenes.

Building names	Architectural pictures
Foguang Temple	
Daming Palace	
Famen Temple	
South Chan Temple Main Hall of Mount Wutai	
Tiantai An Main Hall of Pingshun	
Main Hall of Guangren Temple in Ruicheng	

**Fig. 2** Examples of Chinese Tang Dynasty architecture



Building names	Architectural pictures
Todaiji Temple	
Horyuji Temple	
Tangshodaiji Temple	
Kofukuji Temple	
Yuanxingji Temple	
Ninwaji Temple	

**Fig. 3** Examples of Japanese traditional architecture



Fig. 3 continued

Each augmentation method contributed an additional 1000 images, ultimately forming a comprehensive dataset of 3850 images. The dataset was systematically divided into three subsets: training set (80%, 3080 images), test set (10%, 385 images), and validation set (10%, 385 images). This structured approach ensures a balanced distribution of data during training, validation, and testing of the model, contributing to effective learning and performance evaluation.

**Computer parameter configuration and model training process**

**Hardware setup**

In our experimental environment, we utilized the YOLOv8 model developed by Ultralytics, version 0.114. The environment was configured with Python 3.9.0, VScode (1.76.0) IDE, and CUDA 10.2. All model

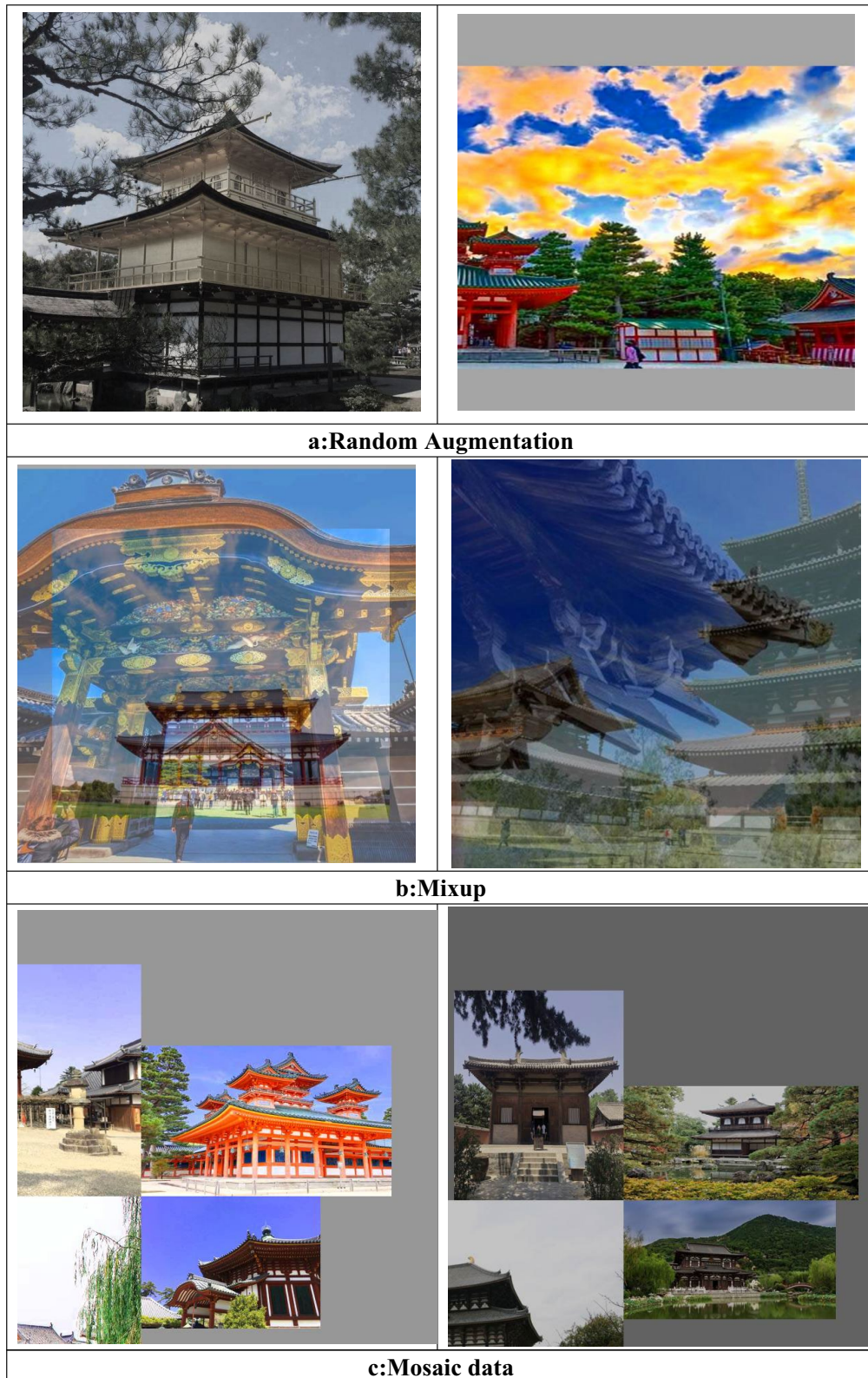
training and testing were performed on an NVIDIA TITAN V (12 GB).

**Model training steps**

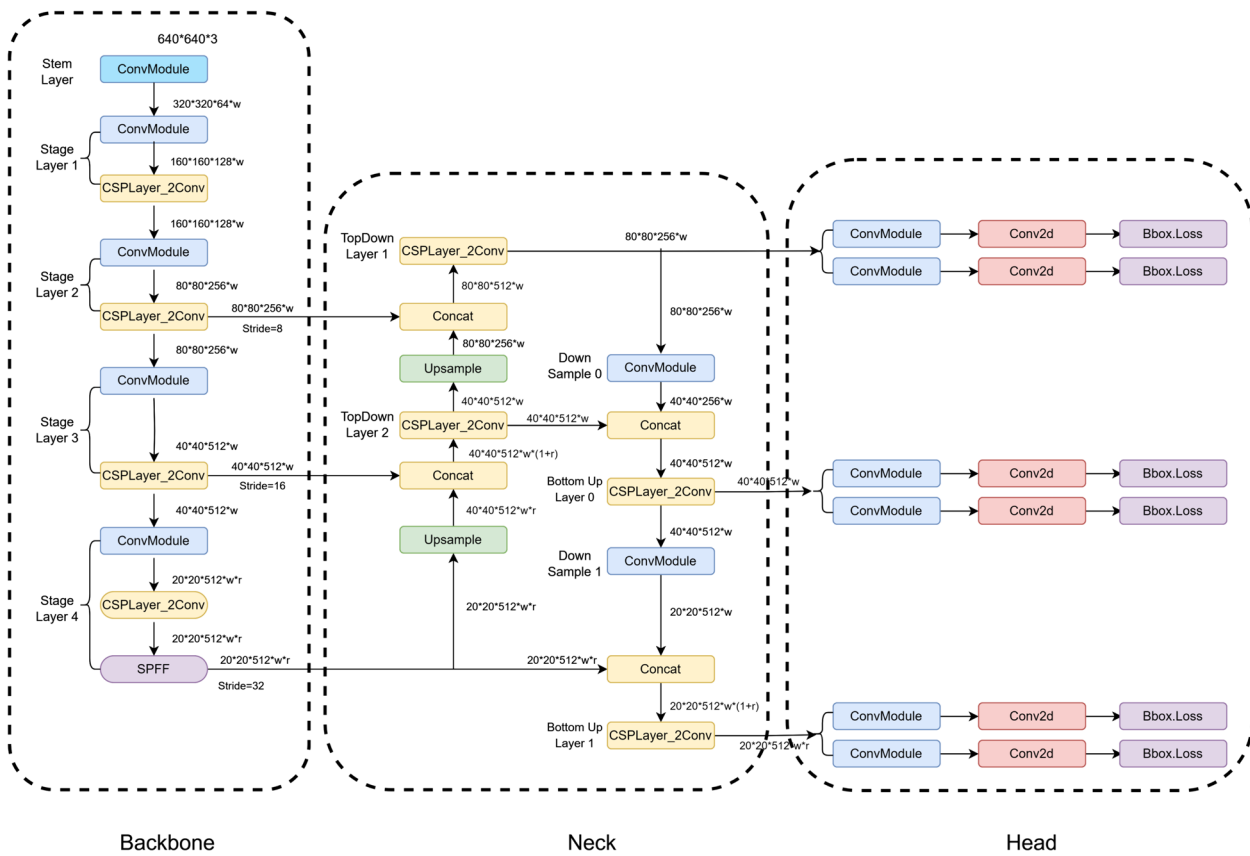
Training parameters were defined as follows: the model was trained for 200 epochs to ensure sufficient time for the learning process. A batch size of 16 balanced computational efficiency and memory utilization. The image size was set to 640×640 pixels, suitable for detailed object detection.

**Architecture details**

The YOLOv8’s backbone has been modified to include MHA blocks, which are integrated at strategic layers within the backbone. This modification allows the model to leverage the strengths of MHA, such as improved handling of spatial relationships across large distances



**Fig. 4** Data augmentation method



**Fig. 5** YOLOv8 Network Architecture Diagram

in input images. The inclusion of MHA in the backbone helps in better recognition of detailed architectural elements by considering broader contextual information that traditional convolutional layers might overlook.

**YOLOv8 model tuning**  
**YOLOv8 model structure**

YOLOv8 (Fig. 5), the latest version in the YOLO series, utilizes a feature pyramid network and cross-layer connections to integrate multi-scale feature information, enhancing the accuracy of object detection. Its core structure includes a custom backbone network (often based on CSPDarknet), which improves gradient flow and reduces computation. YOLOv8 treats object detection as a regression problem and predicts through convolutional layers, pooling layers, and fully connected layers. To enhance real-time detection precision, YOLOv8 meticulously adjusts network parameters and incorporates techniques like anchor boxes, IoU threshold, and non-maximum suppression. It also combines data augmentation and other optimization strategies to enhance model performance [36–39].

**Add MHA module to YOLOv8 model**

Multi-Head Attention (MHA), a pivotal part of the Transformer architecture, enhances its ability to manage long-range dependencies by distributing the attention mechanism across multiple heads (Fig. 6). This approach allows for simultaneous focus on varied features by taking query, key, and value vectors as inputs, conducting weighted summations based on query-key similarities, and combining these with value vectors for output, using dot product or bilinear methods for similarity assessments.

The formal definition of MHA is as follows:

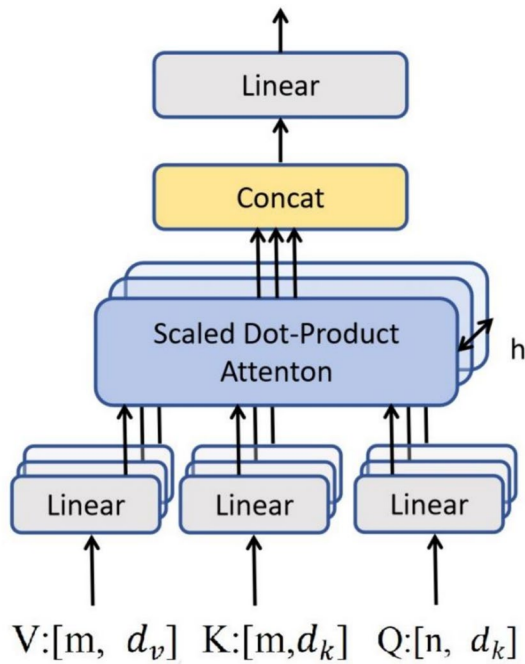
$$MHA = MultiHead(Q, K, V) = Concat(head_1, \dots, head_h) W^o$$

where  $head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$ ,  $i \in 1, 2, \dots, h$ ,  $h$  is the second dimension of the output from the Concat operation is  $h \cdot d_v$ .

The dimensions of each variable are as follows:

$$W_i^Q : [d_{model}, d_k]$$

### Network structure of Multi-Head Attention



**Fig. 6** MHA Network Architecture Diagram

$$W_i^K : [d_{model}, d_k]$$

$$W_i^V : [d_{model}, d_v]$$

The second dimension of the original Q (Query), K (Key), and V (Value) are all the same, and they are projected into corresponding subspaces through linear projection.

$$W^o : [h \cdot d_v, d_{model}]$$

Therefore, MHA can be viewed as the following mapping:

$$MHA : R^{n \cdot d_{model}} \rightarrow R^{n \cdot d_{model}}$$

where n is the length of the sequence, and  $d_{model}$  represents the dimension of each token vector.

SDPA (Fig. 6), standing for Scaled Dot-Product Attention [40], belongs to the multiplicative attention mechanism. In simple terms, it involves weighting the Value based on the degree of match between Query (Q) and Key. In fact, Query, Key, and Value all originate from the input, so fundamentally, SDPA essentially reorganizes the input information.

The formal definition of SDPA is as follows:

$$SDPA = Attention(Q, K, V) = softmax\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \cdot V$$

$V(u_k)$  where the dimensions of each variable are as follows:

Q:[n,  $d_k$ ], representing n Queries, each Query being a vector with a dimension of  $d_k$ ;

K:[m,  $d_k$ ], representing m Keys, each Key being a vector with a dimension of  $d_k$ ;

V:[m,  $d_v$ ], representing m Values, each Value being a vector with a dimension of  $d_v$ ;

Note: K and V appear in the form of Key-Value pairs. Each vector in K corresponds one-to-one with each vector in V.

Therefore, SDPA can be seen as the following mapping:

$$SDPA : (R^{n \times d_k}, R^{m \times d_k}, R^{m \times d_v}) \rightarrow R^{n \times d_v}$$

Note: The input of softmax in the formula is a Matrix, and the softmax operation is performed row by row, that is.

$$softmax(X)_{i,j} = \frac{expX_{i,j}}{\sum_j expX_{i,j}}$$

QKV convolutional blocks generate Query (Q), Key (K), and Value (V) feature maps using specialized convolution layers instead of the typical linear transformations used in standard Transformers. These convolution layers are designed to preserve the spatial integrity of the input images, allowing the attention mechanism to effectively focus on relevant features within the visual field. This method leverages the inherent strengths of CNNs in handling image data, while enhancing the model's capability to attend to and process significant textural and structural information in complex scenes.

#### Model evaluation and performance metrics

For model evaluation, images from the test set are input into the trained model to obtain prediction results. These predictions are then compared with the true labels to calculate various metrics.

In object detection tasks, a predicted bounding box is considered a true positive (TP) if the Intersection over Union (IOU) with a ground truth box exceeds a certain threshold; otherwise, it's considered a false positive (FP). False negatives (FN) are actual objects that the model failed to detect. The study employs several commonly used evaluation metrics in the field of object detection to assess the performance of the algorithm model. The primary metrics include Precision (P),

Recall (R), Average Precision (AP), F1 Score, and Mean Average Precision (mAP).

Precision (P): Represents the proportion of correct predictions among all predicted targets. The formulas for precision is as follows:

$$P = \frac{TP}{TP + FP}$$

Recall (R): Denotes the proportion of actual targets correctly predicted by the model.

The formulas for Recall is as follows:

$$R = \frac{TP}{TP + FN}$$

Average Precision (AP): Refers to the area under the Precision-Recall (PR) curve.

F1 Score: Balances both precision and recall, considering their harmonic mean.

The formulas for F1 Score is as follows:

$$F_1 = \frac{2PR}{P + R}$$

Mean Average Precision (mAP): The average of AP across all categories. mAP<sub>0.5</sub> indicates the mAP at an IOU threshold of 0.5, while mAP<sub>0.5:0.95</sub> represents the mean mAP at different IOU thresholds, ranging from 0.5 to 0.95 with an increment of 0.05.

The formulas for mAP is as follows:

$$mAP = \frac{1}{N} \sum_{i=1}^N \int_0^1 P(R) dR$$

## Results

### Performance of YOLOv8-MHA model

#### Confusion matrix

The confusion matrix (Fig. 7a) for the YOLOv8 model demonstrates high classification accuracy for Tang Dynasty (TDy) and Japanese Tang-style (JP) architectural structures, achieving 87% and 91%, respectively. Although the model's performance in distinguishing background elements is suboptimal, the primary focus on the recognition of TDy and JP categories indicates the model's overall excellence.

#### The precision-confidence curve

Figure 7b demonstrates the model's high precision across categories, nearing perfection at a 0.918 confidence score, with TDy and JP curves sharply peaking and stabilizing,

showcasing its effectiveness in accurate high-confidence predictions and architectural classification.

#### The recall-confidence curve

As shown in Fig. 7c, the model correctly identifies all relevant instances of Tang Dynasty (TDy) and Japanese traditional (JP) architecture in the dataset based on the confidence threshold. The model demonstrates a high recall rate for all categories at lower confidence thresholds, with a recall rate of 0.96 at zero confidence threshold, indicating comprehensive detection of relevant instances.

#### The precision-recall

As shown in Fig. 7d, the curve demonstrates the balance between precision and recall for each category and all category combinations. The precision for TDy is 0.917, while the precision for JP is slightly higher at 0.963. The overall performance of all categories generates a mean Average Precision (mAP) of 0.940 at an Intersection over Union (IoU) threshold of 0.5, indicating strong object detection capabilities.

#### The F1-confidence curve

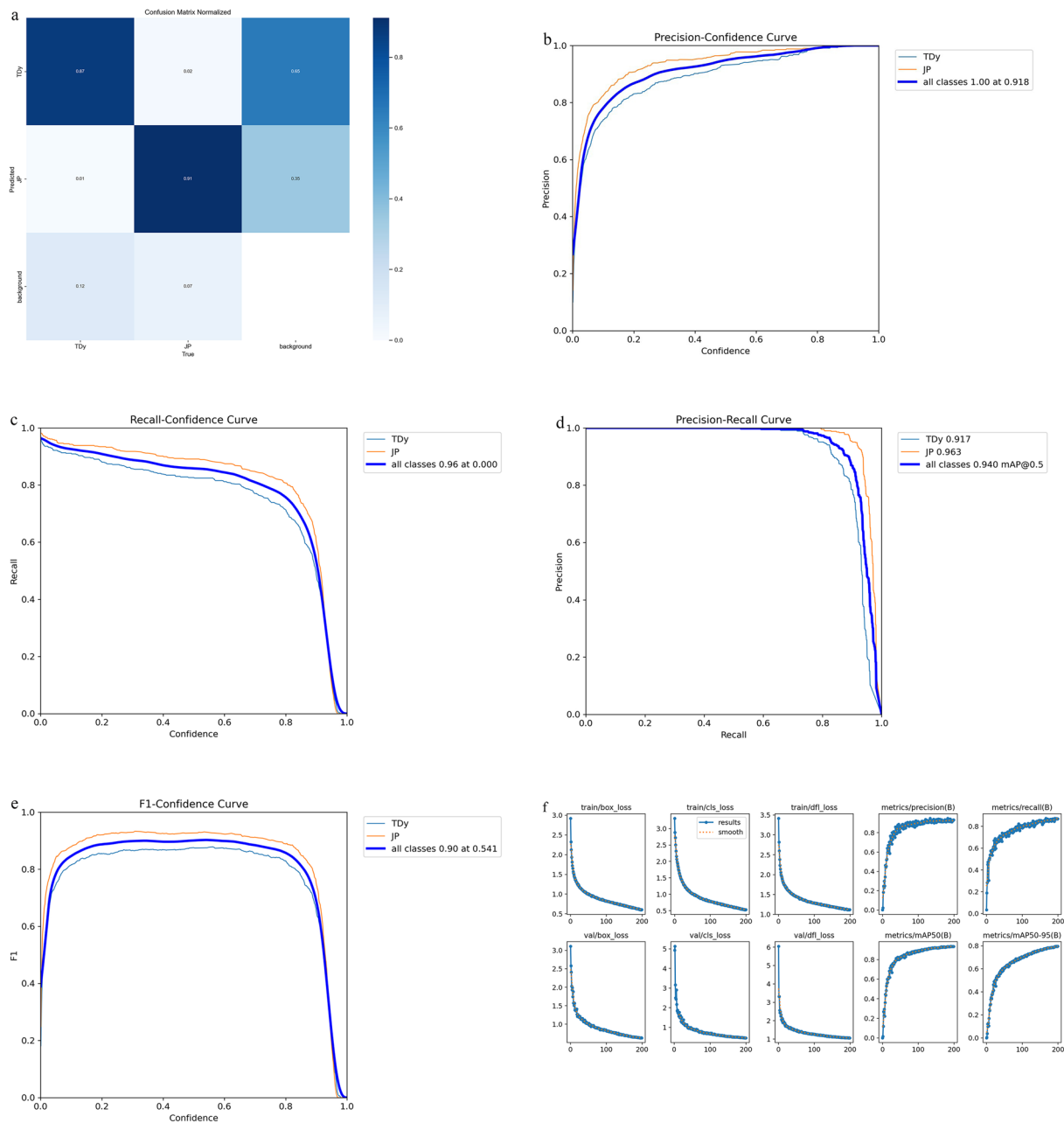
As shown in Fig. 7e, F1-Confidence Curve for TDy and JP architectures reveals both categories attain high F1 scores, showing balanced performance. The model sustains near-peak F1 scores across varied confidence levels, with a notable 0.90 F1 score at 0.541 confidence, underscoring its precision and recall balance, crucial for reliable architectural style differentiation.

### Comprehensive performance analysis (Fig. 7f)

#### Loss graphs

- Training Loss (box, cls, and obj): All three components of training loss—box loss, class loss, and objectness loss—show a sharp decline initially, then gradually decrease, indicating effective learning. The smoothing line indicates steady convergence without significant fluctuations, suggesting a well-tuned learning rate and robust training process.
- Validation Loss (box, cls, and obj): The corresponding validation losses reflect the trends seen in the training losses, indicating that the model generalizes well to unseen data. The absence of an upward trend in later periods suggests that overfitting is unlikely to occur.

#### Performance metrics

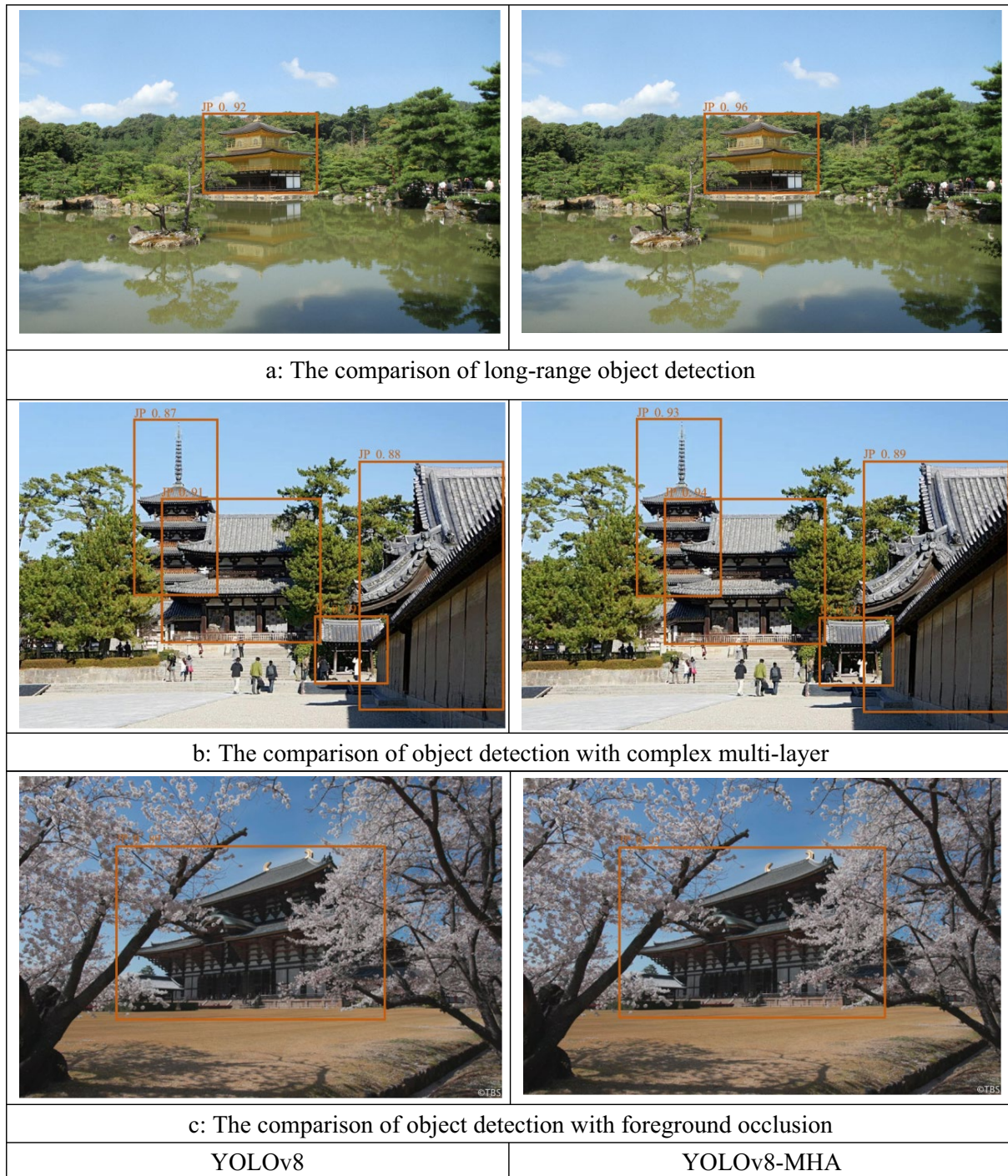


**Fig. 7** The Performance of YOLOv8-MHA model

- Precision and Recall (B): Both precision and recall for the validation set show initial rapid improvement, stabilizing at high values. This indicates that the model can correctly identify and classify objects of interest with high precision and covers most relevant instances in the dataset.
- mAP50 and mAP50-95 (B): Mean Average Precision at IoU=0.50 (mAP50) and the mean value for IoU threshold ranges from 0.50 to 0.95 (mAP50-95) both

show excellent performance, with mAP50 stabilizing near 1. These metrics indicate that the model has highly accurate target detection capabilities within the IoU threshold range, and mAP50-95 reflects robustness under more stringent matching criteria.

Overall, the model demonstrates strong convergence characteristics and high performance on both seen (training) and unseen (validation) data. High precision,



**Fig. 8** The performance differences between YOLOv8 and YOLOv8-MHA

recall rates, and mAP scores across a range of IoU thresholds emphasize the model's reliability and accuracy in object detection tasks.

**Comparative analysis of detection accuracy for Tang dynasty architecture and Japanese traditional architecture**  
 Figure 8 shows the performance of YOLOv8 and YOLOv8-MHA under different environments.



**Table 1** Ablation experiment

Modules	Precision	Recall	mAP@50	mAP@50–95
YOLOv5	0.856	0.825	0.875	0.755
YOLOv8	0.897	0.876	0.905	0.795
YOLOv8-MHA	0.956	0.856	0.94	0.804

### Ablation experiment

This ablation study, presented in Table 1, assesses YOLOv5, YOLOv8, and YOLOv8-MHA models. The YOLOv8-MHA model, with its integration of Multi-Head Attention, surpasses predecessors in precision (95.6%) and mAP scores (mAP@50: 94%, mAP@50–95: 80.4%), establishing its superiority in object detection accuracy.

### Discussion and limitations

#### Discussion

This study's use of the YOLOv8-MHA algorithm offers nuanced insights into Tang Dynasty and Japanese traditional timber structures, facilitating architectural preservation and public education on these traditions. It enhances awareness and conservation of cultural heritage.

YOLOv8-MHA provides significant advancements in detecting and classifying architectural elements, aiding professionals in architecture and cultural history. It brings a deeper understanding and inspiration for integrating historical elements into modern designs, ensuring cultural continuity.

The research faces challenges due to the limited data from surviving Tang Dynasty structures, highlighting the need for ongoing conservation efforts to enrich future studies and improve model generalization capabilities.

#### Weaknesses and limitations

While the YOLOv8-MHA model demonstrates substantial improvements in detecting architectural features, several limitations merit further discussion.

- (1) **Resource Intensity:** The integration of the Multi-Head Attention (MHA) mechanism increases the computational complexity of the model, leading to higher resource consumption during training and inference. This may limit the model's deployment in resource-constrained environments.
- (2) **Generalizability Issues:** Although our model performs well on the datasets tested, its ability to generalize to vastly different architectural styles or degraded historical structures has not been exten-

sively validated. This could affect the model's applicability in broader preservation tasks.

- (3) **Sensitivity to Training Data:** The performance of the YOLOv8-MHA model is highly dependent on the quality and diversity of the training data. Inadequacies in dataset representation can lead to biases in model predictions, particularly for underrepresented architectural styles.

### Proposed improvements

- (1) **Efficiency Optimization:** Future work could focus on optimizing the MHA mechanism to reduce the model's resource demands without compromising its performance. Techniques such as pruning, quantization, and knowledge distillation may be explored to achieve a more efficient model [41–43].
- (2) **Robust Generalization:** Enhancing the model's robustness by training on a more diverse set of images, including those from different periods and in various states of preservation, could improve its generalizability. Employing techniques like domain adaptation might also be beneficial [44–46].
- (3) **Data Augmentation:** Implementing advanced data augmentation techniques that simulate a wide range of architectural deterioration could help in training the model to recognize features in less-than-ideal conditions, thereby enhancing its practical utility in real-world applications [47, 48].

### Conclusions and future work

#### Contributions and implications

This study has successfully demonstrated the efficacy of integrating the Multi-Head Attention (MHA) mechanism into the YOLOv8 model, which we termed YOLOv8-MHA. Our enhanced model has shown substantial improvements in the precision and accuracy of detecting and classifying complex architectural features in heritage structures, achieving a precision of 95.6%, a recall of 85.6%, and a mean Average Precision (mAP@50) of 94%.

- (1) A specialized dataset was constructed. By employing three innovative data augmentation techniques—Random Augmentation, Mixup, and Mosaic—the image dataset of Tang Dynasty architecture and Japanese traditional architecture was successfully expanded from an initial 850 images to 3850 images, significantly enhancing the dataset's diversity and complexity. This rich dataset, divided into training (3080 images), testing (385 images), and validation (385 images) sets, provided comprehensive learning and evaluation opportunities for

the model. These strategies significantly improved the model's robustness and generalization ability, providing robust support for identifying and analyzing Tang Dynasty timber structures and Japanese traditional timber structures architectural styles.

- (2) The enhancement of the YOLOv8 network structure, particularly through the addition of the Multi-Head Attention (MHA) mechanism, improved the model's precision in discerning architectural details. The YOLOv8-MHA model exhibited exceptional performance in handling complex architectural features, achieving a precision of 95.6%, a recall of 85.6%, and an mAP50 of 94%, significantly outperforming its predecessors.
- (3) Experimental results confirmed the outstanding performance of the YOLOv8-MHA model in distinguishing between timber structures of Tang Dynasty and Japanese traditional architecture. With an mAP50-95 reaching 80.4%, the model demonstrated its ability to maintain high detection precision across various IoU thresholds. This performance underscores the potential of YOLOv8-MHA in assisting architectural historians, cultural heritage conservationists, and enthusiasts in identifying, documenting, and preserving heritage buildings.
- (4) Technological Advancement: The integration of MHA into YOLOv8 represents a significant step forward in the use of deep learning technologies for cultural heritage preservation. By refining our model to better capture fine-grained details and contextual relationships within architectural imagery, we aim to contribute to the ongoing improvement of accuracy in this field.
- (5) Practical Application: The YOLOv8-MHA model's increased accuracy and precision are critical for practitioners in the field of heritage conservation, providing them with a more reliable tool for documenting and analyzing historical structures, potentially aiding in restoration and maintenance efforts.

#### Future work

Future research will broaden the application of high-precision architectural style recognition technology to a wider array of cultural heritages, such as sculptures, murals, and manuscripts, aiming for deeper preservation and understanding of diverse cultural values. Efforts will focus on enhancing the model's cross-cultural recognition capabilities and environmental adaptability, developing a dynamic, interactive heritage database, and promoting interdisciplinary collaboration across fields like computer science and digital humanities. Additionally, sustainability in cultural heritage preservation will

be prioritized, utilizing advanced image recognition for effective monitoring and maintenance, thereby ensuring the sustainable protection of invaluable cultural assets.

#### Acknowledgements

Authors would like to acknowledge the financial support of the following funds: Chinese Ministry of Education Humanities and Social Sciences Research Youth Fund Project (Grant numbers: 23YJC760101). The National Natural Science Foundation of China (Grant numbers: 62003137). Zhejiang Provincial Natural Science Foundation (Grant numbers: LTGS23F030003). And authors would like to acknowledge the financial support of Huzhou Key Laboratory of Intelligent Sensing and Optimal Control for Industrial Systems (Grant 2022-17).

#### Author contributions

Author contributions C.G: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing—Original, Draft Visualization. G. Z: Conceptualization, Methodology, Software, Writing—Original, Draft Visualization. S.G: Conceptualization, Methodology, Software, Draft Visualization. S.D: Conceptualization, Methodology, Software, Supervision. E.K and T. S: Methodology, Software, Resources, Writing—Review and Editing, Supervision. All authors have read and agreed to the published version of the manuscript.

#### Data availability

Data will be made available on request.

#### Declarations

#### Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 11 April 2024 Accepted: 15 June 2024

Published online: 28 June 2024

#### References

1. Cui YX. The impact of education and culture during the Sui and Tang dynasties on Japan. *Da Guan*. 2018;02:92–3.
2. Liu LT, Tian RC. On the absorption and inheritance of Chang'an capital culture by Nara Heijo-kyo in Japan. *Humanities Collection*. 2019;01:269–90.
3. Guo Q. A study on the patterns of ridge decoration and dissemination in Chinese and Japanese Tang Dynasty architecture (Master's thesis, Huazhong University of Science and Technology). 2021. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202301&filename=1021907504.nh>.
4. Yang JJ. Patterns of scroll grass in the Asuka and Nara periods of Japan. *Decoration*. 1998;01:49–51. <https://doi.org/10.16272/j.cnki.cn11-1392/j.1998.01.010>.
5. Yuan HG, Liu LN. A brief analysis of the influence of Chinese Tang dynasty architectural art on Japanese traditional architecture. *Architect Cult*. 2018;11:123–4.
6. Zheng LX. The inspiration of Japanese architecture on modern Chinese architectural design concepts. *Edu Inf Technol Forum*. 2018;11:50–1.
7. Li SY. Archaeological study and restoration discussion of the West Ming temple site in Tang Chang'an (Doctoral dissertation, Nanjing University). 2018. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFDLAST2021&filename=1019022393.nh>.
8. Wang Z, Hua Z, Wen Y, Zhang S, Xu X, Song H. E-YOLO: recognition of estrus cow based on improved YOLOv8n model. *Expert Syst Appl*. 2024;238: 122212.
9. Guan H, Deng H, Ma X, Zhang T, Zhang Y, Zhu T, Lu Y. A corn canopy organs detection method based on improved DBI-YOLOv8 network. *Eur J Agron*. 2024;154: 127076.

10. Xiong C, Zayed T, Abdelkader EM. A novel YOLOv8-GAM-Wise-IoU model for automated detection of bridge surface cracks. *Constr Build Mater*. 2024;414: 135025.
11. Cao Y, Pang D, Zhao Q, Yan Y, Jiang Y, Tian C, Li J. Improved YOLOv8-GD deep learning model for defect detection in electroluminescence images of solar photovoltaic modules. *Eng Appl Artif Intell*. 2024;131: 107866.
12. Zhang Y, Zhang H, Huang Q, Han Y, Zhao M. DsP-YOLO: an anchor-free network with DsPAN for small object detection of multiscale defects. *Expert Syst Appl*. 2024;241: 122669.
13. Luo WL. A study on the structural characteristics of the Great Ming Palace during the zenith of Tang dynasty architectural art. *Lan Tai World*. 2014;15:9–10. <https://doi.org/10.16565/j.cnki.1006-7744.2014.15.005>.
14. He XR. The inheritance and development of Tang Dynasty architectural colors in Xi'an urban color scheme (Master's thesis, Shaanxi Normal University). 2021 <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202202&filename=1021139645.nh>.
15. Wang JX. A discussion on the characteristics of Tang dynasty architecture during the peak of China's feudal society. *Time Educ*. 2017;02:141.
16. Yin MJ. A study on the roof decorations of Tang dynasty buildings. *Popular Literature Art*. 2017;21:50.
17. Guo XN. Interpreting the architectural art characteristics of the Tang dynasty from the East Great Hall of Foguang temple. *Lan Tai World*. 2014;06:147–8. <https://doi.org/10.16565/j.cnki.1006-7744.2014.06.003>.
18. Liu YZ. Cultural heritage protection in Japan as seen from the historical relics of Nara and Kyoto. *China Cultural Heritage*. 2010;06:106–10.
19. Shao JZ, Hu ZY. A study on the architectural style development of traditional Japanese wooden towers. *Architect J*. 2015;12:23–9.
20. Yin N. The development and evolution of Japanese Buddhist art: taking the Asuka, Nara, and Heian periods as examples. *Fa Yin*. 2015;02:52–5. <https://doi.org/10.16805/j.cnki.11-1671/b.2015.02.015>.
21. Wan N. A study on ancient Japanese architecture. *Ancient Architecture Gardens Technol*. 2008;02:22–4.
22. Liu LT, Tian RC. The absorption of Tang dynasty Chang'an Buddhist architectural culture in the Nara period of Japan. *Jiangnan Forum*. 2018;06:119–27.
23. Shao ZY, Cao MJ. The influence of Chang'an Qinglong Temple on Nara's Saidai-ji and Genko-ji. *Western Academic Journal*. 2016;07:44–7. <https://doi.org/10.16721/j.cnki.cn61-1487/c.2016.07.013>.
24. Li H. A comparative study of the ratio of eaves height to eaves projection in Chinese and Japanese wooden structure architecture. *Proceedings of the 5th International Symposium on Chinese Architectural History*. 2010; 5: 370–377.
25. Guo JQ. The connection between ancient Japanese temple architecture tile roofing and the roofing practices of China's Tang dynasty. *Ancient Architecture and Gardens Technology*, 1997: 04. CNKI:SUN:GJYL.0.1997-04-005.
26. Chen Z. The influence of Chinese culture and art on the style of ancient Japanese gardens. *Chinese Gardens*. 1986;04:36–9.
27. Liang GM. Research on Japanese temple architecture and its gardens under the influence of Southern Song Zen culture (Master's thesis, Shandong University). 2019. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201902&filename=1019081629.nh>.
28. Ma J, Yan W, Liu G, Xing S, Niu S, Wei T. Complex texture contour feature extraction of cracks in timber structures of ancient architecture based on YOLO algorithm. *Adv Civil Eng*. 2022;2022(1):7879302.
29. Yiming F, Xianxin G, Kun C, Zhu Z, Qing Ye. Accurate and automated detection of surface knots on sawn timbers using YOLO-V5 model. *BioResources*. 2021. <https://doi.org/10.15376/BIORES.16.3.5390-5406>.
30. Kurdthongmee W, Suwannarat K. Locating wood pith in a wood stem cross sectional image using YOLO object detection. 2019; <https://doi.org/10.1109/TAAL48200.2019.8959823>
31. Bayomi N, El Kholy M, Fernandez JE, Velipasalar S, Rakha T. Building envelope object detection using YOLO models. 2022. <https://doi.org/10.23919/ANNSIM5834.2022.9859463>
32. Cui Z. Automatic classification of ancient architectural components based on deep learning (Master's thesis, Beijing University of Civil Engineering and Architecture). 2021. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202201&filename=1021634516.nh>.
33. Ling Q. Classification research of ancient architectural cultural heritage protection knowledge (Master's thesis, Chinese Academy of Sciences). 2008. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2008&filename=2008131524.nh>
34. Shi ZL. Types and classification of ancient Chinese architecture (Master's thesis, Zhejiang University). 2004. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD9904&filename=2004075896.nh>
35. Llamas J, Leronés PM, Medina R, Zalama E, Gómez-García-Bermejo J. Classification of architectural heritage images using deep learning techniques. *Appl Sci*. 2017;7(10):992. <https://doi.org/10.3390/app7100992>.
36. Reis D, Kupec J, Hong J, Daoudi A. Real-time flying object detection with YOLOv8. *arXiv preprint arXiv:2305.09972*. 2023.
37. Zou MY, Yu JJ, Lv Y, Lu B, Chi WZ, Sun LN. A novel day-to-night obstacle detection method for excavators based on image enhancement and multi-sensor fusion. *IEEE Sens J*. 2023;23:10825–35.
38. Wang N, Liu H, Li Y, Zhou W, Ding M. Segmentation and phenotype calculation of rapeseed pods based on YOLO v8 and mask R-convolution neural networks. *Plants*. 2023;12(18):3328.
39. Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 21–26 July 2017 (pp. 6517–6525). 2017.
40. Lou H, Duan X, Guo J, Liu H, Gu J, Bi L, Chen H. DC-YOLOv8: small-size object detection algorithm based on camera sensor. *Electronics*. 2023;12(10):2323.
41. Li J, Wang X, Tu Z, Lyu MR. On the diversity of multi-head attention. *Neurocomputing*. 2021;454:14–24.
42. Chen H, Jiang D, Sahli H. Transformer encoder with multi-modal multi-head attention for continuous affect recognition. *IEEE Trans Multimedia*. 2020;23:4171–83.
43. Lin Y, Wang C, Song H, Li Y. Multi-head self-attention transformation networks for aspect-based sentiment analysis. *IEEE Access*. 2021;9:8762–70.
44. Du Y, Pei B, Zhao X, Ji J. Deep scaled dot-product attention based domain adaptation model for biomedical question answering. *Methods*. 2020;173:69–74.
45. Jiang B, Chen S, Wang B, Luo B. MGLNN: Semi-supervised learning via multiple graph cooperative learning neural networks. *Neural Netw*. 2022;153:204–14.
46. Roy AM, Bhaduri J. DenseSPH-YOLOv5: an automated damage detection model based on DenseNet and Swin-Transformer prediction head-enabled YOLOv5 with attention mechanism. *Adv Eng Inform*. 2023;56: 102007.
47. Roy AM, Bhaduri J, Kumar T, Raj K. WilDect-YOLO: an efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection. *Eco Inform*. 2023;75: 101919.
48. Qiao W, Zhao Y, Xu Y, Lei Y, Wang Y, Yu S, Li H. Deep learning-based pixel-level rock fragment recognition during tunnel excavation using instance segmentation model. *Tunn Undergr Space Technol*. 2021;115: 104072.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.