

RESEARCH

Open Access



# Detection and recognition of Chinese porcelain inlay images of traditional Lingnan architectural decoration based on YOLOv4 technology

Yanyu Li<sup>1†</sup>, Mingyi Zhao<sup>1†</sup>, Jingyi Mao<sup>2</sup>, Yile Chen<sup>1\*</sup> , Liang Zheng<sup>1\*</sup>  and Lina Yan<sup>1,3</sup> 

## Abstract

With the rapid development of machine learning technology, it has become possible to automatically identify cultural heritage elements in traditional buildings. This research aimed to develop a machine learning model based on the YOLOv4 architecture to identify the traditional Chinese porcelain inlay pattern in the Lingnan region. The researchers collected and annotated a large quantity of Lingnan Chinese porcelain inlay image data and then used these data to train the studied model. The research results show that (1) the model in this study was specifically adjusted to effectively identify a variety of Chinese porcelain inlay pattern types, including traditional patterns such as plum blossoms and camellias. (2) In the 116th epoch, the model showed excellent generalization ability, and the verification loss reached the lowest value of 0.88. The lowest training loss in the 195th epoch was 0.99, indicating that the model reached an optimal balance point for both recognition accuracy and processing speed. (3) By comparing different models for detecting Chinese porcelain inlay images across 581 pictures, our YOLOv4 model demonstrated greater accuracy in most classification tasks than did the YOLOv8 model, especially in the classification of chrysanthemums, where it achieved an accuracy rate of 87.5%, significantly outperforming YOLOv8 by 58.82%. However, the study also revealed that under certain conditions, such as detecting apples and pears in low-light environments, YOLOv8 showed a lower missing data rate, highlighting the limitations of our model in dealing with complex detection conditions.

**Keywords** Machine learning, YOLOv4 model, Traditional Lingnan architectural patterns, Cultural heritage element identification, Chinese porcelain inlay

<sup>†</sup>Yanyu Li and Mingyi Zhao have contributed equally to this work and share first authorship.

\*Correspondence:

Yile Chen

2009853GAT30001@student.must.edu.mo

Liang Zheng

2009853GAT30002@student.must.edu.mo

<sup>1</sup> Faculty of Humanities and Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa 999078, Macau, China

<sup>2</sup> College of Arts, South China Agricultural University, No.483 Wushan Road, Tianhe District, Guangzhou 510642, China

<sup>3</sup> Planning Department, Shanghai Total Architectural Design & Urban Planning Co, Ltd, Shanghai, China

## Introduction

### Research background: the importance of the chinese porcelain inlay in the inheritance of lingnan culture

Chinese porcelain inlay is a unique ethnic architectural decoration technology in the Lingnan region of China. It originated during the Wanli period of the Ming Dynasty (1573–1620) and flourished in the Qing Dynasty and the Republic of China. It was included in China's second batch of national intangible cultural heritage lists in 2008 [1] and China's third batch of national intangible cultural heritage expansion projects in 2011 [2]. It is one of the symbolic representatives of Lingnan culture and ancestral

hall culture. Using materials first uses gray molding as the base. Then, according to the requirements of various forms, special scissors, grinding wheels, rough pliers, and other tools are used to make corresponding porcelain pieces from various colors of thin and crisp low-temperature glaze porcelain (such as bowls and plates). It is also used for surface decoration, and various images are embedded and decorated on the building surface. Therefore, Chinese porcelain inlay works are vivid and durable without deformation or fading [3]. Chinese porcelain inlay craft is an important part of Lingnan history and culture. Numerous factors, including a strong foundation in local folk culture, have influenced and shaped its emergence and development. It is perfectly integrated with ancient Yue culture, Central Plains culture, and foreign culture and is highly individual and creative, making it a wonder in the field of traditional Chinese architectural decoration [4]. As early as the Ming and Qing Dynasties, Chinese porcelain inlay was very popular in the Lingnan area due to its bright colors, exaggerated shapes, exquisite skills, diverse themes, and magnificence. Princes, nobles, wealthy gentries, merchants, literati, and ordinary people flock to the Chinese porcelain inlay craft in traditional architectural decorations and are proud to use Chinese porcelain inlay craftsmanship to decorate their family ancestral halls. The traditional architectural decorative craft of Chinese porcelain inlays can not only meet aesthetic needs for traditional architectural art but also reflect the need for the spirit of traditional Chinese culture [5]. In today's era, with the rise of artificial intelligence technology, if it can be combined with traditional architectural decorative images such as Chinese porcelain inlays, it is a topic worth exploring.

## Literature review

### *Chinese porcelain inlay art and culturally related research*

In traditional Cantonese-style architecture, straw ash is used to lay the prototype, and after drying, paper ash is used to shape the detailed building facade. Colored gray sculpture is one of the unique outdoor decoration techniques in this area. For example, the Guangzhou Chen Clan Ancestral Hall (1888–1894) is a typical representative work of this technique. Some scholars believe that due to the hot and rainy climate in the Chaoshan area, its proximity to the ocean, and the high salt content of the air, traditional gray sculpture art in the Lingnan area is not durable when used as a decorative method for local buildings. Chaoshan craftsmen drew inspiration from ceramics and combined them with the original gray sculpture art form in Guangdong to create a new craft technology called Chinese porcelain inlay. While retaining the decorative charm and value, it also considers corrosion resistance and durability. This tradition has been

passed down to this day [6]. In related research on Chinese porcelain inlay technology, some scholars have analyzed the visual and symbolic significance of the artistic expression of Chinese porcelain inlays. They believe that Teochew porcelain inlay, whether it is the location chosen for decoration on the edge of the building or the patterns and decorative trends adopted, best embodies the traditional Chinese aesthetic concept of “flying beauty” and carries the spiritual symbols and sustenance of the Teochew people [7]. In addition, few scholars have studied the Chinese porcelain inlay outside the Chaoshan area. Some scholars have explored the creative transformation and innovative development of traditional architecture in Liwan, Guangzhou, and the Lingnan area. The original materials and processes of Chinese porcelain inlay have been mentioned along with assumptions on how to use new materials to express the charm of traditional buildings containing Chinese porcelain inlay components [8]. When studying the Temple of the Dawn in Thailand, some scholars have suggested that the artistic technique of floral decoration porcelain components in the temple tower probably came from the Chaoshan region of China. They believe that “during the period of the Bangkok Dynasty, as the political and commercial exchanges between China and Thailand increased, a large number of Chinese people migrated from the Chaoshan area to Thailand, and porcelain inlay technology was introduced to Thailand” [9]. Despite this, at this stage, research on Chinese porcelain inlays has focused mainly on the sources of craftsmanship, craftsmanship techniques, craftsmanship characteristics, and the content and significance of decorative themes.

### *Application of artificial intelligence technology in cultural heritage and art*

With the rapid development of artificial intelligence technology in recent years, AI technologies such as YOLO and GNNs [10] have been widely used in: monitoring and protecting wild endangered animals [11], identifying and preventing plant diseases and insect pests [12], identifying and monitoring civilian damage to infrastructure [13], and other fields. The problem-solving methods and abilities of this technology are also applicable to the fields of art and cultural heritage, thus triggering additional thinking and research. For example, the YOLO series of models can be used to quickly identify and detect target objects in cultural heritage [14–20]. These studies are very important for the regular inspection and maintenance of architectural heritage. In addition, some scholars have conducted image recognition research on the decorative construction of mythical beasts on the roofs of traditional Chinese buildings based on the YOLOv3 model [20]. Similarly, some researchers have also used

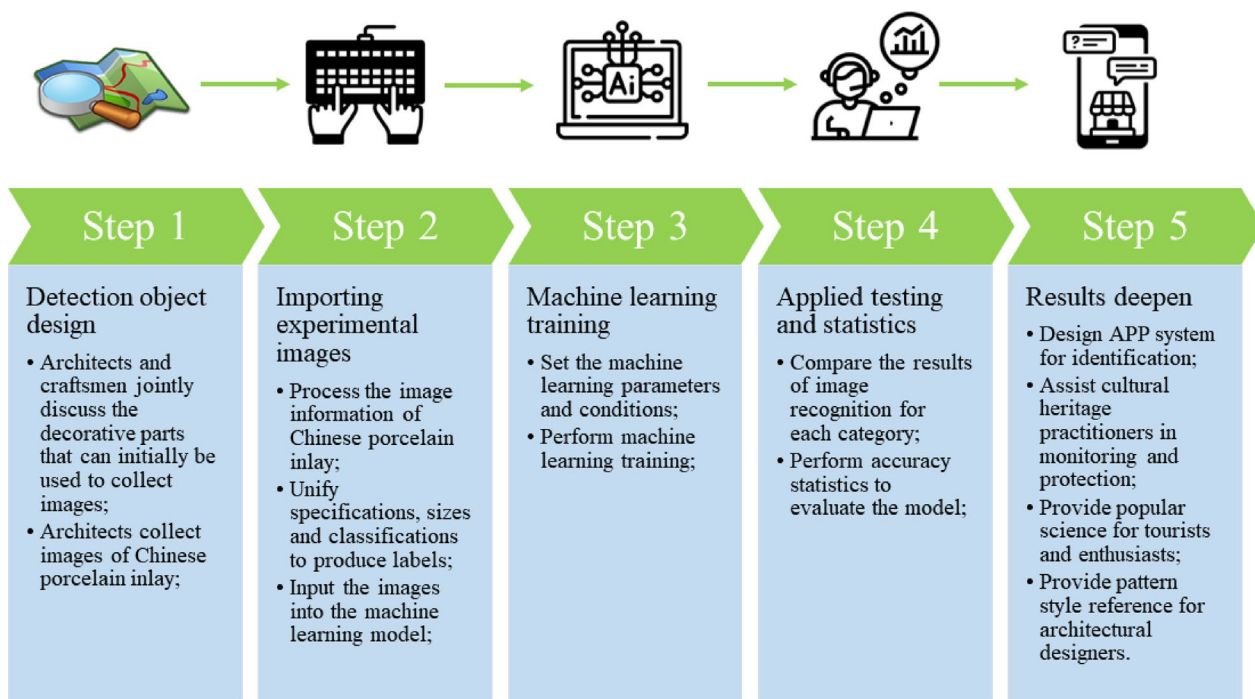
the object detection method built by the YOLO model to conduct classification detection experiments on building types in specific areas of Athens, Greece, since 1834 to quickly determine the artistic style of buildings [21]. In addition to the YOLO series of technologies, other artificial intelligence models have made progress in the image recognition of architectural or cultural heritage [22]. For example, specific cases include religious architectural sites [23], monuments [24], China’s famous historical residences, Hakka Weilong Houses [25], architectural heritage in China’s Hubei region [26], Indonesian batik craft patterns [27], rock art patterns [28], and Chinese national costume images [29]. Based on this, researchers are encouraged to see that artificial intelligence technology is playing an increasingly critical role in the cultural heritage and art field, and its applicable scenarios and scope are gradually deepening with the accumulation of research. However, it is undeniable that the current research and discussion in this field are still not extensive or sufficient. In particular, for cultural heritage sites with obvious regional characteristics, there is still much room for research.

**Problem statement and objectives**

When appreciating the works of Chinese porcelain inlays, people often focus on wonderful characters, auspicious animals, and other themes for the first time but ignore the most numerous and least important life themes in

Chinese porcelain inlays. This type of Chinese porcelain inlay often plays the role of foil, but it is an indispensable part of Chinese porcelain inlay work. The works with life themes are very diverse, and among them, the works with plant themes are prolific. Coupled with the artistic creation and personalized expression of craftsmen, people often cannot accurately identify patterns with life themes [30]. People always make guesses about plant patterns based on personal experience and cannot immediately find the corresponding intangible inheritors or craftsmen for verification. This creates a bias in the understanding of the cultural connotation of images, and sometimes, incorrect perceptions will remain throughout their lives. In recent years, with the continuous development of machine learning and algorithms, computer vision technology has gradually been applied in various fields. In terms of cultural heritage protection, the trend toward intelligence and information technology has become increasingly obvious. At present, the pattern style of Chinese porcelain inlays is mainly manually recognized, which is highly subjective and uncertain.

If machine learning technology can automatically identify and locate the types and locations of the intangible cultural heritage of Chinese porcelain inlay images, it will greatly help tourists, traditional architecture enthusiasts, experts, and scholars quickly understand the specific information of this traditional architectural decoration (as shown conceptually in Fig. 1). This technology can



**Fig. 1** Concept of artificial intelligence for assisting in Chinese porcelain inlay identification

also be used as a strong support for the artistic 3D reconstruction of Chinese porcelain inlay on the roofs of traditional Lingnan buildings, which is conducive to the digital protection and inheritance of cultural heritage. It can also further strengthen the role of the Chinese porcelain inlay in cultural exchanges and dissemination. Therefore, it is highly important to carry out research on the detection and recognition of Lingnan traditional Chinese porcelain inlay patterns based on machine learning.

However, due to different aesthetic preferences and decorative purposes, Chinese porcelain inlay patterns have many different scales and complex arrays, posing challenges in accurate identification. These patterns are often intertwined, and some elements cover other elements; thus, it is difficult to distinguish all their elements. In addition, the patterns of some themes are richer than those of other themes, forming an unbalanced information database in which fewer emerging themes are not easy to identify. The multiscale, mutual occlusion and data imbalance of these patterns reflect common problems in dense object positioning and prediction. The YOLOv4 model is famous for its efficient positioning and prediction of dense objects. The application of this technology in Chinese porcelain inlay recognition can provide a reference solution to the above problems.

Based on the abovementioned limitations of traditional methods for identifying Chinese porcelain inlay pattern styles and the advantages of YOLOv4, this article addresses the following three main research questions:

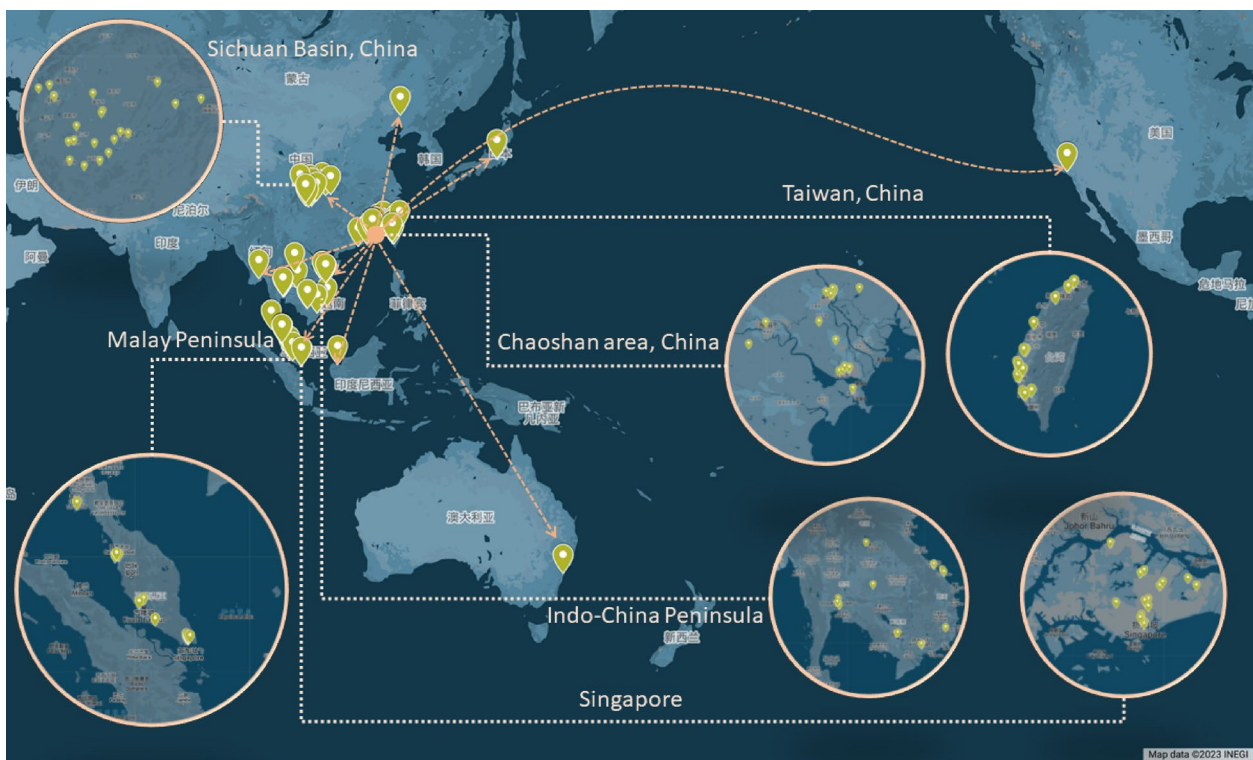
- (1) Which categories of Chinese porcelain inlay pattern styles can be used as training sets for training and identification?
- (2) How is the specific technical process constructed under the YOLOv4-based model?
- (3) What is the result of Chinese porcelain inlay application recognition?

**Materials and methods**

**Analysis of chinese porcelain inlays in traditional lingnan architecture**

*Characteristics of chinese porcelain inlay*

The charm of the Chinese porcelain inlay is not only reflected in its “fish-scale bird wings” and “drama-like” architectural decoration but also, more importantly, as a carrier of traditional Chinese culture [27]. More importantly, it is regarded as a disseminator of traditional Chinese culture (Fig. 2). In today’s multipolar world, an increasing number of “Chinese elements” are appearing on the world stage, forming a new trend of regional



**Fig. 2** The main road map for the spread of Chinese porcelain inlay (image source: drawn by the author; the base map of the map comes from google maps)

spread around the world. As an architectural decoration technology that combines tradition and modernity, the Chinese porcelain inlay is representative of China's excellent traditional culture and has considerable room for expansion in promoting international cultural communication [31]. In addition to the Fujian and Chaoshan areas of Guangdong in China, there are also many traditional Chinese buildings overseas that use Chinese porcelain inlay for decoration (Fig. 3). It is also widely distributed in Malaysia, Singapore, Vietnam, Thailand, Myanmar, Japan and other countries. Chinese porcelain inlay appears not only in traditional religious temples but also in former Chinese residences, chambers of commerce, and even in Chinese cemeteries and worship-type buildings (please refer to Appendix A for the specific distribution of buildings).

Chinese porcelain inlay finished products can be divided into three main types: flat inlay, floating inlay, and three-dimensional inlay (Fig. 4). Flat inlays are generally used for close-up scenes and small patterns. It is mainly used on the side ridges and pediments of traditional buildings. It is often used in combination with floating inlays. Floating inlays are mostly used on roofs and are usually shaped based on painting patterns to pursue multilayered decorative effects. It is mainly used on the roof ridges and gables of ancestral halls, temples, and other buildings and on the foyers and walls of traditional

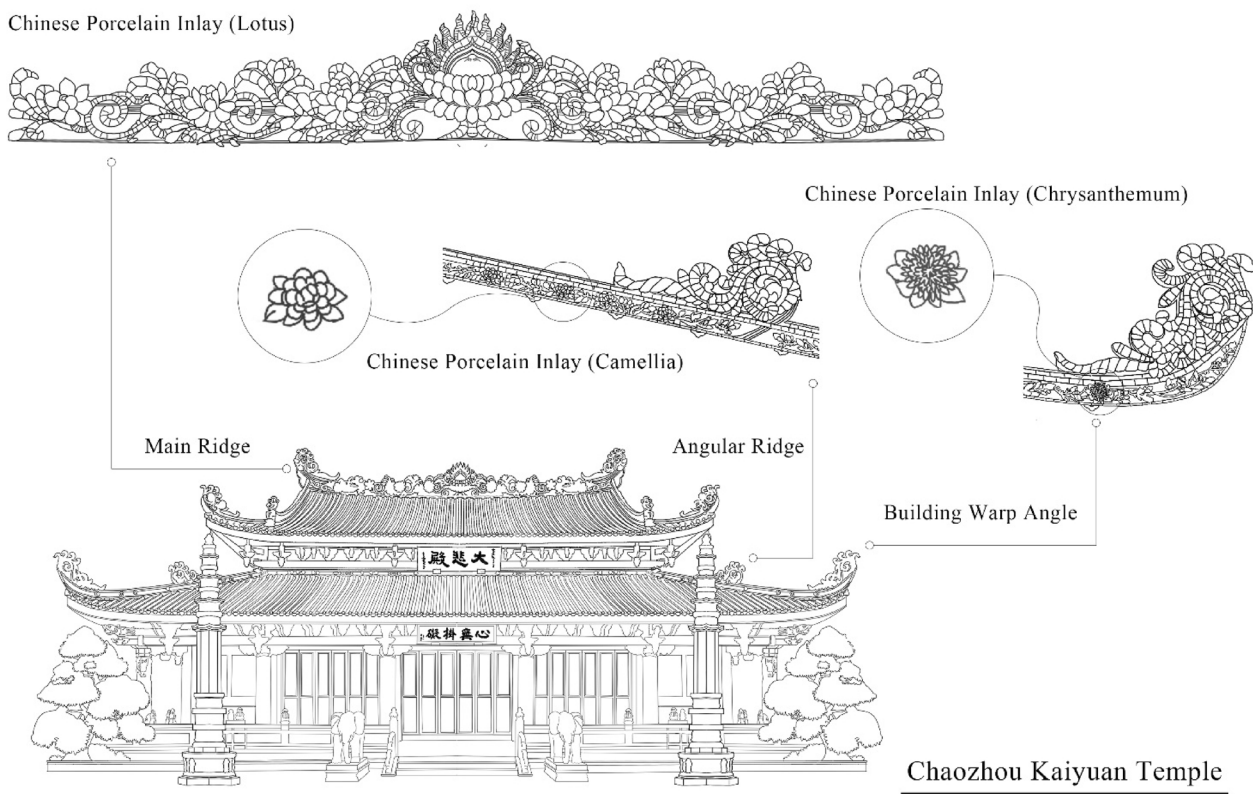
houses. Three-dimensional inlays are the most complex and difficult of the three forms. They focus on the fullness, uniformity, and symmetry of the picture. The final effect can be viewed in 360°. They are mainly used on the main and vertical ridges of buildings.

**Classification of themes and meanings of chinese porcelain inlay**

The themes of Chinese porcelain inlay mostly revolve around peace and auspiciousness, expressing reverence for gods, worship of ancestors, and yearning for a better life. Most of them are based on the characteristics of the subject matter or the homophonic auspicious meanings. For example, shrimp and crabs symbolize "good harvest", chrysanthemums symbolize "long life", magpies symbolize "good news comes," and vases symbolize "peace and success, ". Many themes can be combined to form new meanings. For example, the combination of plum blossoms and magpies symbolizes "brows full of happiness," and the combination of pine, cypress, and white cranes symbolizes "songs and cranes prolong life," expressing a yearning for a better life. During the creation process, craftsmen often perform artistic processing on the corresponding themes, such as thickening the branches, enlarging the flowers, and stacking petals to create a clear, full, prominent, and stable visual effect. Occasionally, they violate the natural laws of plant growth by



**Fig. 3** The main distribution locations of the architectural decorative art Chinese porcelain inlay in Asia (image source: drawn by the author; the base map of the map comes from Google Maps)



**Fig. 4** The style characteristics of Chinese porcelain inlay and its main distribution position in traditional Lingnan architecture

displaying flowers from different seasons in the same picture, arranging plum blossom branches with narcissus leaves, and designing banyan leaves in colorful colors. The Chinese porcelain inlay symbolism of common plant and flower themes is shown in Table 1 below.

**Image data and research process**

**Image data source**

In this study, the main images collected by the researchers came from the roofs of traditional Lingnan buildings, as shown in Table 2. According to the researchers' early investigations, there is currently no ready-made dataset that can be used for image-style detection of the intangible cultural heritage of Chinese porcelain inlays. Therefore, the researchers collected a thousand first- and second-hand images of Chinese porcelain inlays through three in-depth field surveys and typing keywords on the internet. The first-hand Chinese porcelain inlay images for this article were collected in Chaozhou City and Shantou City, Guangdong Province, China (Table 2). These include Kaiyuan Temple, Qinglong Ancient Temple, Guanyin Temple, Cunxin Shantang, the Inlaid Porcelain Museum, and the Daliao Inlaid Porcelain Crafts Society. The image acquisition equipment used was a SONYα6400. The size of the collected first- and

second-hand images ranges from 325×437 pixels to 4457×6685 pixels, and the percentage of Chinese porcelain inlay-style images in the total area of the images ranges from 1 to 90%, as shown in Fig. 3. According to previous research, the training effect is better when the number of images is 100–200 times the number of labels. Therefore, in this study, there are seven main Chinese porcelain inlay-style labels, which means that 700–1400 Chinese porcelain inlay-style images are needed for machine learning model training. However, due to the disrepair of some Chinese porcelain inlay works, the porcelain pieces fell off and were seriously damaged, which greatly affected the recognition effect of the image. The researchers screened more than 400 Chinese porcelain inlay-style images, leaving 707 Chinese porcelain inlay-style images as experimental samples. These samples can provide sufficient information for the model to understand the characteristics and rules of the Chinese porcelain inlay style.

**Research process**






In this article, researchers propose a method based on the YOLOv4 model that can automatically identify specific pattern names from a large number of Chinese porcelain inlay image styles. This not only provides a

**Table 1** The Chinese porcelain inlay symbolism of common plant and flower themes

Plant theme	Plant characteristics	Features of Chinese porcelain inlay production	Meaning	Photos
Peony	The flowers are large with bright colors	The petals are stacked in multiple layers with irregular edges, and the flowers have exaggerated proportions. A single flower has more than two colors, often appearing in pieces	Among the flowers, peony ranks first and is known as the "King of Flowers". Liu Yuxi wrote, "Only peonies are the true national beauty, and they move the capital when they bloom". They symbolize wealth, good luck, prosperity, and national beauty	
Camellia	The leaves are green and bright, the flowers are luxuriant, dignified and elegant	The flowers are larger and more regular, rich in color, with various petals in various shapes, regular edges, and arranged in a spiral around the center point	Camellia is one of the ten most famous flowers in China. Gui Youguang praised it as "Although it has a rich and noble appearance, it is not voluptuous. It will wither in the cold years, and it will be like the spring breeze". It symbolizes humility, virtue, and prudence	
Chrysanthemum	There are many types, each with a long flowering period, bright flowers, and alternating leaves	There are various varieties of chrysanthemums in Chinese porcelain inlays. The porcelain inlay is tapered and has a gathering shape in the center	Chrysanthemums are known as "hermits among flowers" because of Tao Yuanming's poem "Pick chrysanthemums under the eastern fence, leisurely see the Nan Mountains", which means good luck and longevity, rapid success, high moral integrity, and integrity. Combined with plum, orchid, and bamboo, it is called the "four gentlemen" among flowers	
Plum blossom	The bark is light gray or greenish, and the flowers are fragrant and bloom before the leaves	The flowers are small, spiral around the stamens, and appear in clusters. The flowers are visible, but not the leaves. Chinese porcelain inlay craftsmen sometimes paste some leaves to enrich the picture	"The Book of Songs" says, "There are beautiful flowers in the mountains, and plum blossoms and plum blossoms are waiting for them." The plum blossom is the first of the "Four Gentlemen" among the flowers. Together with the pine and bamboo, it is also known as the "Three Friends of the Winter". It stands proud in the snow and frost, which symbolizes the character of a scholar, elegance and purity, five blessings, good news, and peace	
Lotus	The stems are long and thick, the leaves are round, and the petal ends are slightly pointed or blunt	The flowers are large and have many gradient colors. They are generally used as the main theme of Chinese porcelain inlay decoration	The lotus is a symbol of Buddhism. It is said that "the flower blooms to see the Buddha's nature," symbolizing the four virtues of Buddha's Dharma: happiness, self, and purity. Zhou Dunyi wrote a poem describing that it "comes out of the mud but is not stained, washes the ripples cleanly but is not evil". The lotus is beautiful but not vulgar, symbolizing holiness, wealth, and good luck	

Source: author statistics

**Table 2** Statistical location of image collection during fieldwork






Acquisition time	Name & location	Picture	Collection position	Quantity
14/03/2021 13/01/2022 05/03/2022	Chaozhou porcelain inlay museum		The exhibition hall	20
15/03/2021	Lu's ancestral Hall		The exterior facades of all buildings	30
15/03/2021	Congxi temple		The exterior facades of all buildings	6
16/03/2021	Three-mountain king ancestral temple		The exterior facades of all buildings	10
16/03/2021	Qinglong ancient temple & Anji King temple		The worship pavilion; main hall; Immortal Master's Palace; official hall	130



**Table 2** (continued)

Acquisition time	Name & location	Picture	Collection position	Quantity
16/03/2021	Kaiyuan temple		Four Heavenly Kings Hall; Heaven King Hall; Mahavira Hall; the scripture perusal chamber; Guanyin Pavilion; the Sixth Patriarch Huineng Hall; Kṣitigarbha Pavilion	130
17/03/2021	Chaozhou West lake		Ao yu	2
17/03/2021 05/03/2022	Chaozhou Tianhou palace		The exterior facades of all buildings	50
08/03/2022	Daliao village, Shantou city		Xu's Ancestral Hall Daliao Porcelain Inlay Craft society	200
08/03/2022	Cunxin Benevolence		The exterior facades of all buildings	60

**Table 2** (continued)

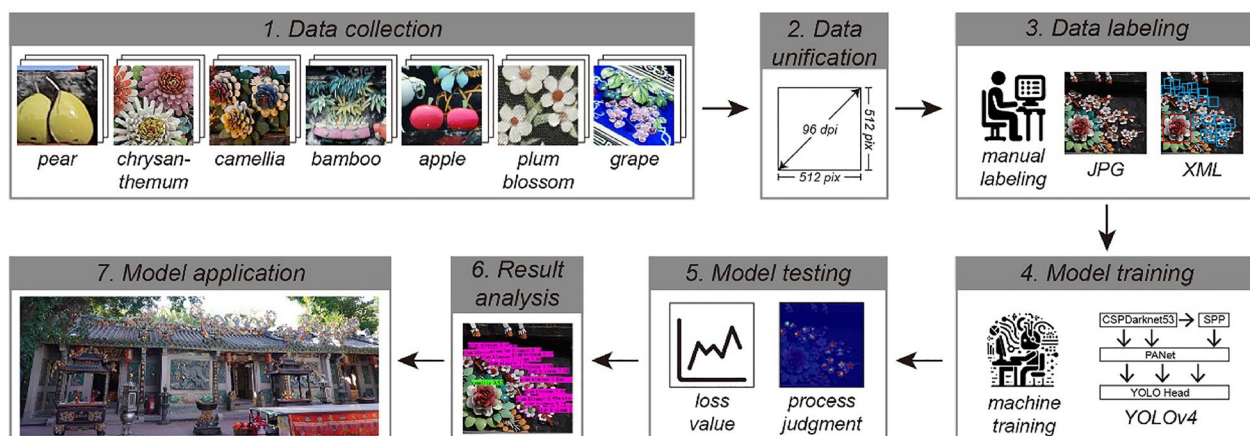
Acquisition time	Name & location	Picture	Collection position	Quantity
08/03/2022	Shantou Little park		Xu Shaoxiang Porcelain Embedding studio	5
09/03/2022	Goddess of Mercy temple in Shantou city		Dafeng temple Goddess of Mercy temple Nine-Dragon screen	45
09/03/2022	Shantou Waima road third primary school		The exterior facades of all buildings	5
09/03/2022	Shantou Tianhou palace		The exterior facades of all buildings	20
09/03/2022	Shantou Cultural Center		The exhibition hall	5

Source: author statistics

scientific and efficient method for the digitization and protection of cultural heritage but also provides a new perspective for its application in modern architecture and design. The main processes included the collection of Chinese porcelain inlay image data, data processing,

data annotation, model training, model testing, and result analysis (Fig. 5).

- (1) Data collection: there is a clear correlation between the stability and generalization ability of model per-



**Fig. 5** Research Process

formance and the diversity and representativeness of the data. Chinese porcelain inlay image data collection and annotation are mainly realized through a large number of Chinese porcelain inlay-style image datasets collected and annotated manually by the author during fieldwork and on the internet. Therefore, during the data collection stage, particular emphasis is placed on gathering data that are both highly representative of Chinese porcelain inlay images and totally diversified in terms of how they are represented. This research team acquired 707 high-definition photos of Chinese porcelain inlay patterns that contained a wide variety of plants and flowers with allegorical themes. These photographs were collected in numerous representative places in the Lingnan region. A variety of weather conditions, including bright, overcast, and rainy days, as well as diverse light situations, including sunlight, shadow, and artificial light sources, were utilized during the image-gathering process. This was done to ensure that as many actual scene changes as possible were captured. The data collection process for Chinese porcelain inlay images was carried out several times from 2021 to 2022 to capture the different performances of Chinese porcelain inlay in different environments. To a certain extent, multiple dimensions of the image are ensured.

(2) Data processing: in regard to the data processing step, a number of image preprocessing procedures are utilized to optimize the quality of the images and guarantee the consistency of the data input. This ultimately results in an improvement in the effectiveness and efficiency of the model training process. Histogram equalization processing

and noise-filtering technology are two examples of these types of technologies. The objective of this approach is to minimize the influence of environmental elements such as light and shadows in the image, reduce the random noise in the image, and improve the clarity of features that have been damaged (Fig. 6). In addition, image size standardization is performed to guarantee that all of the images that are entered into the model have the same resolution and size. This means that all of the images are altered to have a resolution of  $512 \times 512$  pixels, 96 dpi horizontal and vertical resolution, and 24-bit depth. Through the implementation of this action, the consistency of the model training is guaranteed and computational complexity that arises as a consequence of the inconsistent sizes and proportions of the images decrease. Finally, the code for data improvement operations, such as rotating and flipping photos, is also introduced to the program portion of the training process for the model. This is done to broaden the scope of the dataset, improve the model’s capacity for generalization, and strengthen its robustness.

(3) Data annotation: providing precise and consistent annotation information for each image is the primary emphasis of the research team. This is because the data annotation stage is an essential step in the process of ensuring that the model acquires accurate features. Data annotation is performed to prevent errors that are caused by the varying subjective standards of different individuals. Despite the fact that the theme styles of Chinese porcelain inlays have been properly identified, there are still some disparities in the assessments of different people during the actual data annota-



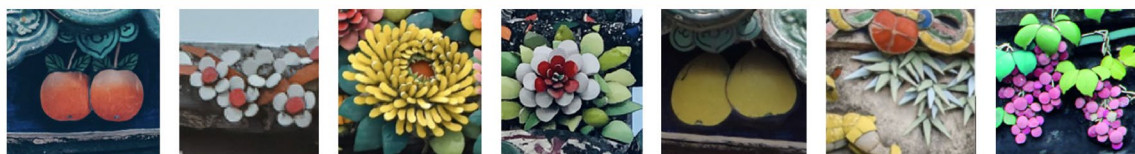
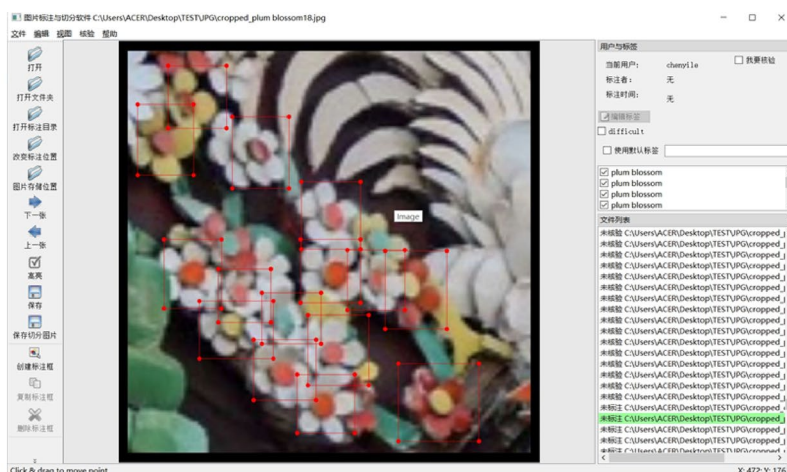
**Fig. 6** Some photos of the Chinese porcelain inlay

tion process. This is based on previous experience. Consequently, to carry out all of the data annotation work, individuals who had been working in the linked field for a considerable amount of time and who were well-versed in their knowledge of plants and flowers were selected. The professional image annotation program LABELIMG is utilized to build precise bounding boxes for the Chinese porcelain inlay pattern styles that are present in each image [32] and performed under the direction of specified annotation criteria. To establish an accurate correlation, a one-of-a-kind code is provided for each type of Chinese porcelain inlay pattern style based on its name. There are seven main labels: chrysanthemum, plum blossom, camellia, bamboo, grape, apple, and pear (Fig. 7). Among them, plum blossoms and camellias had the largest number of samples, with 227 and 259 pictures, respectively, while bamboo, apple, and pear had the least number of samples, with 39, 32, and 20 pictures, respectively. During this study, the researchers were divided into two groups, one of which classified the Chinese porcelain inlay images one by one after consulting botanical professionals and relevant Chinese porcelain inlay craftsmen. The preprocessed Chinese porcelain inlay images are labeled into catego-

ries. Another team double-check the first labels to ensure that the data are accurate. During the process of annotation, this study strongly emphasized the consistency and accuracy of the annotations. To guarantee that each annotation box and category label are accurate, two rounds of annotation review and calibration are performed.

- (4) Model training: in this study, the YOLOv4 object detection framework is used to train the model to achieve accurate recognition of Lingnan Chinese porcelain inlay images. First, the model parameters are initialized, and the input size and anchor box size of YOLOv4 are adjusted according to the characteristics of the Chinese porcelain inlay image to ensure that the model can adapt to Chinese porcelain inlay styles of different shapes and sizes. The training process uses the stochastic gradient descent method with momentum (SGD with momentum). The initial learning rate is set to 0.001, and the momentum is 0.9. In addition, applying weight decay and learning rate decay techniques can effectively avoid overfitting. Model training lasts for a total of 200 training cycles. The training process uses a cross-entropy loss function combined with category loss and localization loss to optimize the model. After each training epoch,

LABELIMG software label production interface



Label name: Apple Plum blossom Chrysanthemum Camellia Pear Bamboo Grape

**Fig. 7** Pattern style name of Chinese porcelain inlay when making labels

the model performance can be evaluated using an independent validation set, and the model can be tuned and optimized based on the average precision (mAP) and loss values.

(5) Model testing: during the process of testing the model, the primary objective is to evaluate the performance of the model on a dataset that is distinct from the one used for training the model. This process is performed to forecast the impact that the model will have when it is actually deployed. This study took great care in preparing a varied test set, which consisted of a total of 37 photos comprising seven different varieties of Chinese porcelain inlay. The purpose of this study is to thoroughly investigate the model’s capacity for generalization. Several different performance indicators are utilized in the testing of the model in this study. The “average precision” (AP) and “miss rate” (MR) are utilized according to the indications of the algorithm [33, 34]. These two indicators can indicate the accuracy of the model as well as its miss rate. In most cases, the characteristics of the algorithmic indicators do not accurately reflect the real detection capabilities of the model. The final model detection results are personally judged and counted one by one to determine the final model accuracy. This ensures that the quality of the model is accurately reflected in practical applications.

(6) Results analysis: The goal of the results analysis phase of model testing is to provide a comprehensive understanding of the performance of the model as well as the potential areas for development. The purpose of this study is to provide a thorough summary of the model’s performance on the overall test set. Additionally, whether the model can accurately and reliably identify and locate Chinese porcelain inlays of various plant-themed types that are used in architectural ornamentation is investigated. Subsequently, the researchers separately remove the detected target area, extract the local features of the image in the area, and compare the image features with the pretrained model image features to identify the specific style name of the detected image. At the microlevel, the focus of this study is on the detection effect of the model on each plant theme pattern type. Additionally, the benefits and drawbacks of the model in terms of detecting various patterns are investigated. Take, for instance, the question of whether there is a greater sensitivity or divergence to particular pattern themes or particular light environment conditions. This study highlights the issues that the model may face in real-world applications by conducting an in-depth investigation of the model’s performance in a variety of situations and conditions. Additionally, this study provides information on how to better optimize the model to

overcome these challenges. In addition, this study investigates the potential failure scenarios of the model and the possible causes behind those failure scenarios, such as data imbalances and features that are not significant. The purpose of this investigation is to provide inspiration for the following improvements and optimizations.

### **Model parameter settings**

This study made targeted adjustments to the YOLOv4 model to improve its efficiency in identifying Chinese porcelain inlay patterns. In terms of the input layer of the model, the resolution of the input image is uniformly adjusted to  $512 \times 512$  pixels. This adjustment helps to increase the speed of image processing without losing key image details. The convolution kernel size of the feature extractor is changed from the standard  $3 \times 3$  to  $5 \times 5$  because of the complexity and variety of the Chinese porcelain inlay pattern. This process is performed so that more detailed features can be captured. To enhance the adaptability and generalization ability of the model, a variety of data enhancement techniques are applied during the training process. For example, the rotation angle of the image is set between  $-30^\circ$  and  $+30^\circ$ , the zoom ratio is adjusted from 0.8 to 1.2, and the brightness and contrast of the image are also appropriately adjusted. These enhancement techniques help the model adapt to different viewing angles and lighting conditions, thereby improving its robustness in practical applications. In addition, to adapt to different lighting and background conditions in practical applications, researchers have conducted scene-specific adaptive training on the model. These detailed adjustments and optimizations are designed to improve the overall performance of the model in Chinese porcelain inlay recognition tasks while ensuring that it can adapt to changing actual application environments.

## **Results: model training and results analysis**

### **Overview of the model architecture**

#### **Advantages of the YOLOv4 model for chinese porcelain inlay image recognition**

In this study, the researchers employed the YOLOv4 model as the basis for detecting Chinese porcelain inlay images, primarily due to the key architectural differences between it and the latest model, YOLOv8, which affords YOLOv4 unique advantages in handling such specific images. First, the CSPDarkNet53 backbone network of YOLOv4, along with its specific feature fusion strategies, such as spatial pyramid pooling (SPP) and path aggregation network (PAN), provide robust feature extraction

capabilities for complex patterns and detail-rich porcelain inlay images. This is particularly important for accurately identifying subtle pattern details. Second, the anchor-based method adopted by YOLOv4, as opposed to anchor-free detection in YOLOv8, offers more precise bounding box localization, which is crucial for detecting small and closely packed objects common in porcelain inlay images. Moreover, YOLOv4 exhibits better adaptability to low-resolution images by design, which is critical for handling porcelain inlay artwork images that may have deteriorated in quality due to their age. While YOLOv8 has achieved significant overall performance improvements, especially in terms of speed and multi-tasking capabilities, its new backbone architecture and anchor-free detection mechanism might not perform as well as the specifically optimized YOLOv4 model when dealing with specific types of complex and detail-dense images. Additionally, although YOLOv8's new loss function shows excellent performance across various tasks, it may still require targeted adjustments and optimizations for highly specialized tasks such as porcelain inlay image detection. Based on these key differences between the architectures of YOLOv4 and YOLOv8, researchers believe that YOLOv4 is more suitable for detecting Chinese porcelain inlay images. It not only provides the necessary feature extraction capability and precise object localization but also inherently adapts to the specific image types and task requirements focused on in this study.

### **YOLOv4 model architecture design**

- (1) Framework principle: the network architecture and key component configuration of the YOLOv4 model are shown in Fig. 8. The model is based on the CSPDarkNet53 backbone network, which takes images with a size of  $512 \times 512 \times 3$  and uses the Mish activation function and multiscale residual block (ResBlock) structure to make learning and extracting features easier [35]. The design of this network architecture aims to maintain the efficiency of deep neural networks in capturing complex patterns while reducing the number of calculations. The YOLOv4 model architecture employs an intricate design to enhance feature extraction and object detection across various scales. At its core, YOLOv4 utilizes CSPDarkNet53, which is designed for efficient feature extraction with a focus on reducing computational complexity while maintaining high performance as its backbone. This backbone is complemented by two critical structures, the path aggregation network (PANet) and the SPP structure, which are pivotal in optimiz-

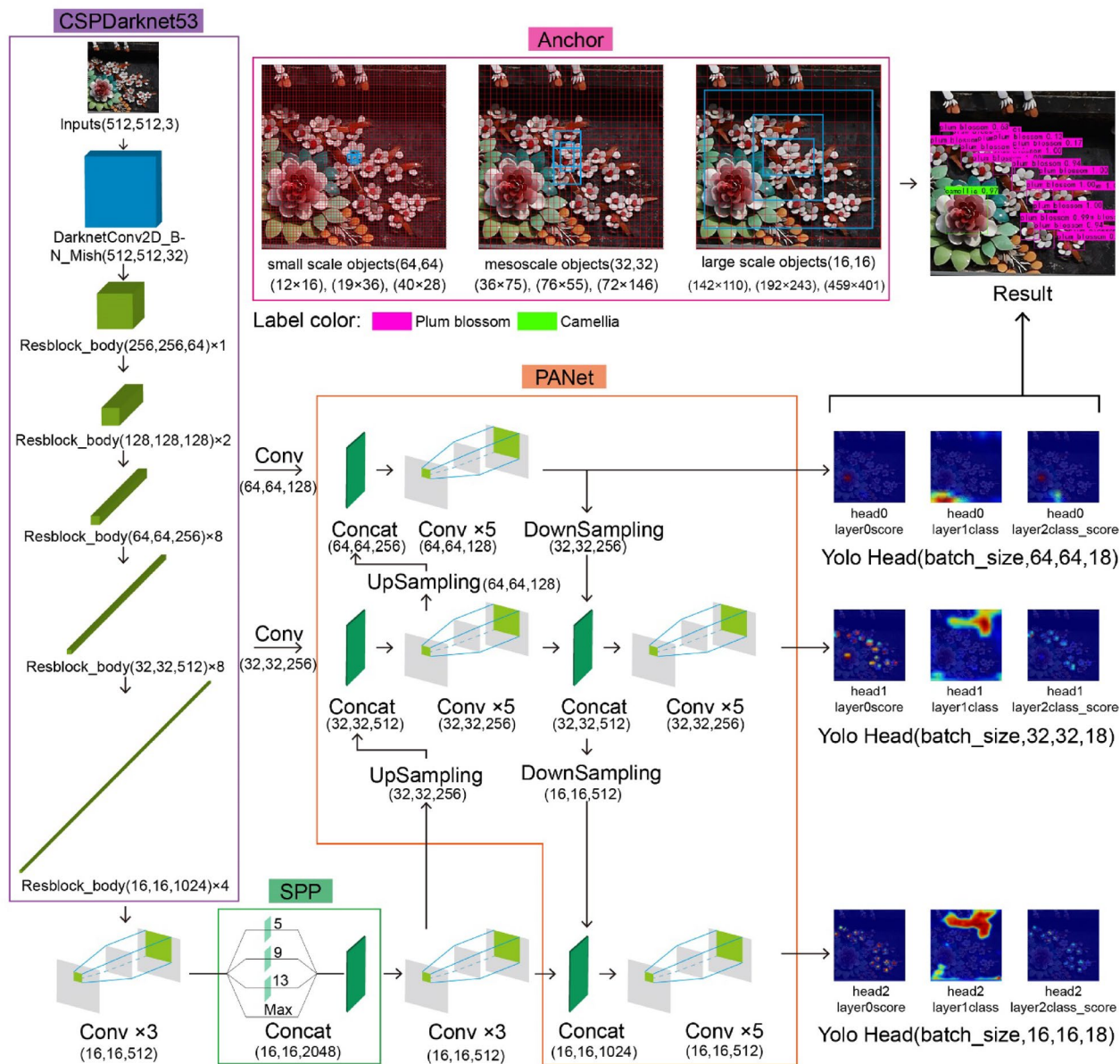


Fig. 8 YOLOv4 model architecture design

ing the model’s performance for multiscale feature fusion.

- (2) PANet structure: the PANet structure plays a vital role in enhancing the feature fusion strategy by improving the flow of information between layers of different resolutions. It achieves this through a strategic process of upsampling and downsampling feature maps, thereby facilitating effective integration of features across scales. This process ensures that both high-level semantic information and low-level detail are accurately captured and utilized,

enhancing object detection across a wide range of sizes.

- (3) SPP structure: in contrast, the SPP structure significantly widens the model’s receptive field by applying spatial pooling over varying scales. By aggregating features under different spatial resolutions, the SPP enables the model to capture a broader context of the input image, thus improving its robustness to variations in object size and appearance. This capability is particularly beneficial for detecting objects

that may significantly vary in scale within the same scene.

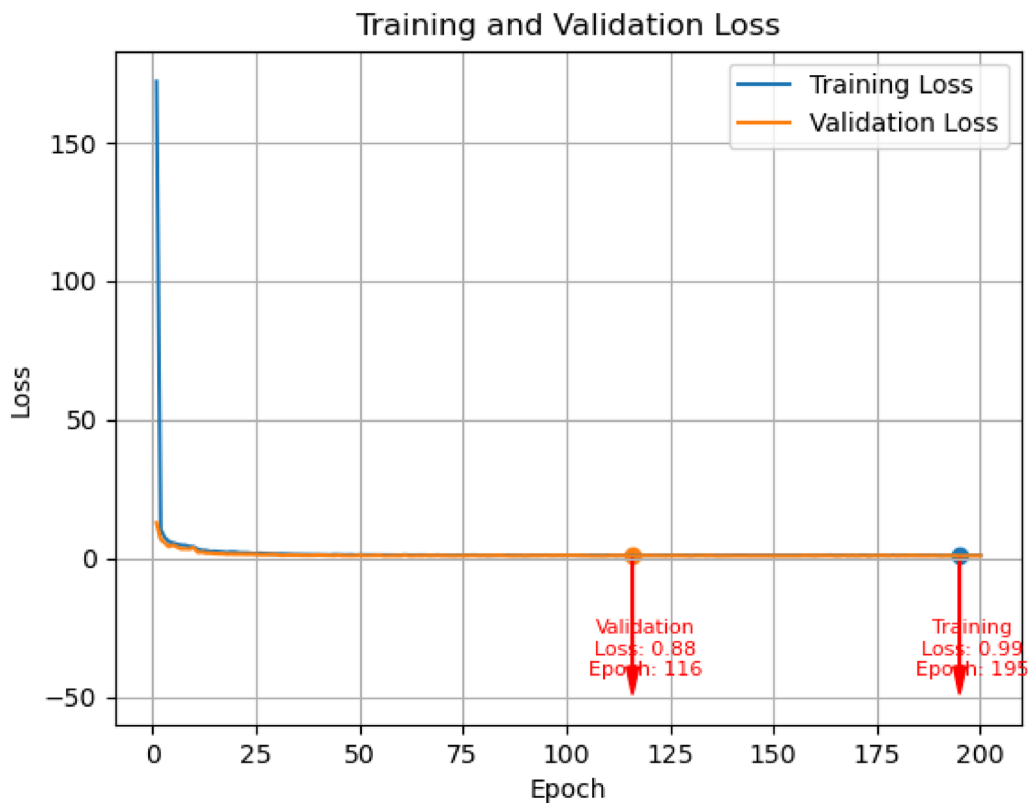
- (4) **YOLO Heads:** to encapsulate the detection process, YOLOv4 incorporates three distinct detection heads (YOLO Heads) at the end of the model, each tailored to process feature maps at different resolutions—specifically,  $64 \times 64$  for large objects,  $32 \times 32$  for medium-sized objects, and  $16 \times 16$  for small objects. This hierarchical approach allows for precise object detection across all levels by efficiently allocating detection tasks based on object size, ensuring that each detection head specializes in capturing objects within a specific scale range.

Furthermore, YOLOv4 introduces several advancements to improve detection accuracy and model efficiency. These include the use of the Mish activation function for nonlinear processing without the drawbacks of traditional ReLU, the integration of cross-stage partial connections (CSPs) to facilitate the learning of more diverse features with less computational demand, and the adoption of anchor-based mechanisms that are meticulously optimized to cover a wide range of object sizes and shapes encountered in various detection scenarios.

Overall, the structural designs and mechanisms employed in YOLOv4, from its CSPDarkNet53 backbone to sophisticated feature fusion and multiscale detection strategies, collectively ensure that the model can effectively handle targets of diverse sizes with high accuracy and speed. This comprehensive approach to model architecture not only addresses the challenges of real-time object detection but also significantly reduces model complexity without compromising performance, making YOLOv4 a highly efficient and versatile solution for object detection tasks.

**Model training**

During the model training process, the training and validation losses were monitored to evaluate the learning effect and generalization ability of the model. As shown in Fig. 9, as the training period (Epoch) increases, the training loss and verification loss of the model decrease significantly, indicating that the model has learned the patterns in the data and can effectively generalize to unseen data. In the early training stages, the loss value drops rapidly, indicating that the model quickly adapts to the training data. After approximately 25 epochs, the loss decreases, and the model begins to converge (refer to Appendix C for loss metrics during training).



**Fig. 9** LOSS trend during model training



Specifically, in the 116th epoch, the verification loss decreased to its lowest value, reaching 0.88. The model at this time showed optimal generalization performance. Although the training loss continued to decrease in the following epochs, finally reaching 0.99 at the 195th epoch, the slight increase in the validation loss indicates that the model may be starting to slightly overfit. This phenomenon shows that after the model learns sufficient generalization rules, more training does not improve the verification performance. Nonetheless, to comprehensively evaluate the performance of the model at different stages, this study selected the 116th epoch, the 195th epoch, and the 200th epoch for further testing.

The reason for selecting these three models is based on their key performance during the training process: the model at the 116th epoch was selected due to having the lowest validation loss, which is usually a criterion for selecting a model's generalization ability. The model of the 195th epoch is based on considering the lowest training loss, which represents the optimal performance of the model on the training set. The model of the 200th epoch represents the final training state, providing a performance snapshot of the model after a complete training epoch. Testing these three models in subsequent sections provides a more accurate assessment of the performance and stability of the models in the practical applications.

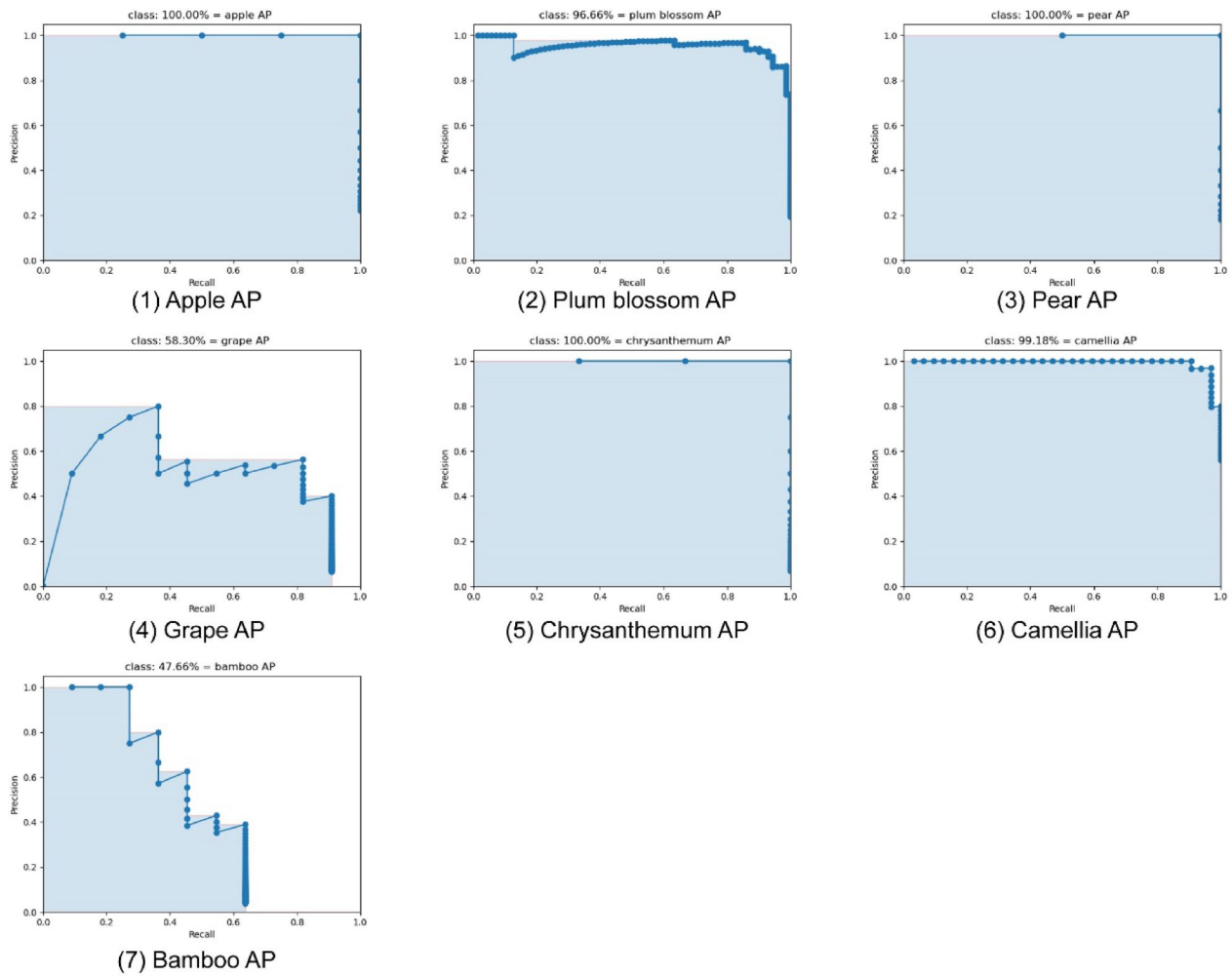
#### Model comprehensive performance evaluation

In the field of machine learning, performance evaluation is a key step to measure the accuracy and effectiveness of the model, especially in visual tasks such as Chinese porcelain inlay category detection. This study uses a range of carefully selected indicators. This includes average precision (AP), F1 score, precision, recall, mean average precision (mAP), and log-average miss rate to comprehensively evaluate the performance of the YOLO model [36, 37]. These indicators comprehensively reflect the model's ability to locate and identify Chinese porcelain inlay patterns. Among them, AP and mAP quantify the accuracy of the model under various confidence thresholds, and the F1 score balances the precision and recall rates, providing a single metric to evaluate the overall performance of the model. The log-average miss rate measures the model's recognition performance for difficult-to-detect objects. Through these measurements, we can gain a deep understanding of the advantages and limitations of the model and guide subsequent model optimization and improvement work. The following is a specific description of the comprehensive performance evaluation of the model in this study.

- (1) It can be observed in the AP plot of the model (Fig. 10) that the detection performance of most

categories is excellent. Due to their distinctive and consistent morphological features, which are well-represented in the dataset and that the model has successfully learned, the Apple and Pear categories achieve 100% AP. The AP of the plum blossom category is slightly lower (96.66%), which may be due to the certain similarity between its morphological characteristics and the background or other Chinese porcelain inlay categories, resulting in a small number of misidentifications. The relatively low AP of the grape category (58.30%) indicates that the model has difficulty identifying its morphology. This may be due to the large variability in the shape of grapes. For example, the colors of grapes include red and purple, and some shapes will be stroked while others will not. In addition, the shapes of grape leaves are also different, and some are artistically processed and do not match the actual situation. Therefore, the model cannot fully learn their characteristics. The high AP of the chrysanthemum and camellia categories (100 and 99.18%) illustrates the model's high accuracy in identifying these categories because these forms have unique and highly distinguishable characteristics. The lower AP of the bamboo category (47.66%) indicates that the features of this category perform inconsistently in the data set. In Chinese porcelain inlay applications, bamboo often appears together with the pattern of the subject, which results in more noise in the graphics and reduces the detection performance of the model.

- (2) In the evaluation of the F1 score, the performance of different Chinese porcelain inlay categories in the model can be observed, and some obvious trends can be observed. As a single indicator, F1 can comprehensively reflect the detection accuracy and completeness of the model. The following is an analysis based on Fig. 11: the F1 scores of the apple and pear categories both reached a perfect score of 1.0 when the score threshold was 0.5. This shows that the model has extremely high precision and recall in these two categories, with almost all real inlays detected and a very low false-positive rate. The F1 curve for plum blossom F1 decreases as the score threshold is lowered because, at lower thresholds, the model starts to incorrectly label more irrelevant regions as plum blossoms, thus increasing the number of false positives. The F1 score of grape (Grape F1) shows large fluctuations, and the highest score is lower than other categories, which may mean the instability of the model in distinguishing grape categories. The F1 scores of chrysanthemum (Chrysanthemum F1) and camellia

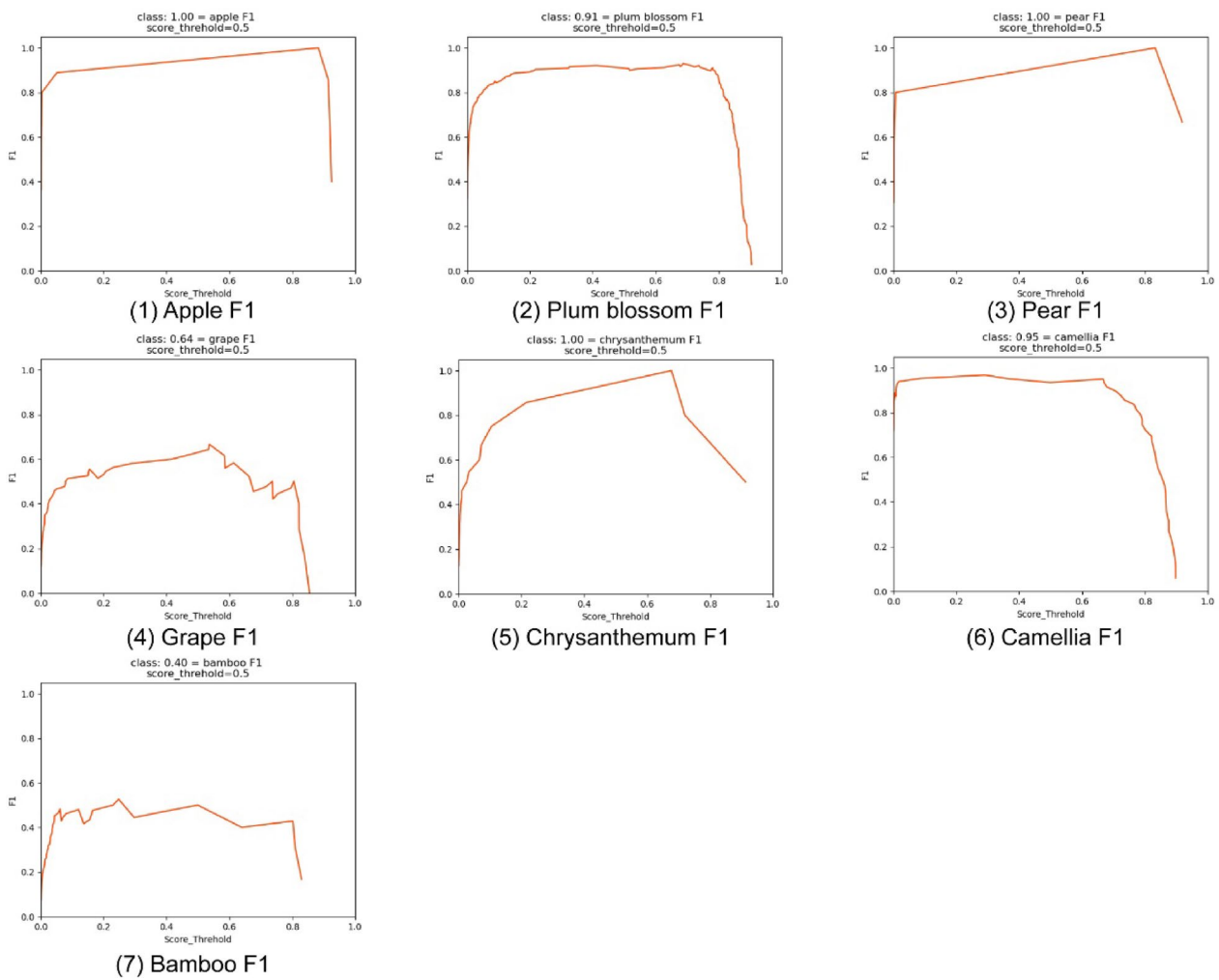


**Fig. 10** Model average precision (AP) statistical analysis

(Camellia F1) are also relatively high, close to 1 and 0.95, respectively, showing that the model has good recognition capabilities in these two categories, with both precision and recall reaching high levels.. Bamboo F1 has the lowest F1 score, which indicates that the model has both a high miss detection rate (low recall rate) and a high false detection rate (low precision) when detecting the bamboo category. The reason is that the morphological diversity and complex background of bamboo make it difficult to distinguish the models.

- (3) Precision is a key indicator for evaluating model performance. It measures the proportion of samples identified as positive by the model that are actually positive. High accuracy means a lower false-positive rate. As shown in Fig. 12, the following analysis of the detection performance of each category can be performed: the accuracy of the apple and pear categories remains at 100%. This shows that

the model has almost no misjudgments in the recognition of these two categories of objects; that is, almost no objects outside these categories are mislabeled. The accuracy of plum blossom decreases slightly when the score threshold is increased. This may mean that under a stricter score threshold, the model’s prediction of plum blossoms becomes more conservative, resulting in real plum blossoms sometimes not being detected; the accuracy of grapes fluctuates greatly, and the overall trend increases when the score threshold increases. This fluctuation may reflect the inconsistency of grape morphology in the data set, causing the model to perform unstably at certain score thresholds. The accuracy of chrysanthemums and camellias is close to 100%, indicating that the prediction of this category is accurate and there are very few false positives; the accuracy of bamboo is the lowest and shows certain fluctuations when the score threshold changes. This



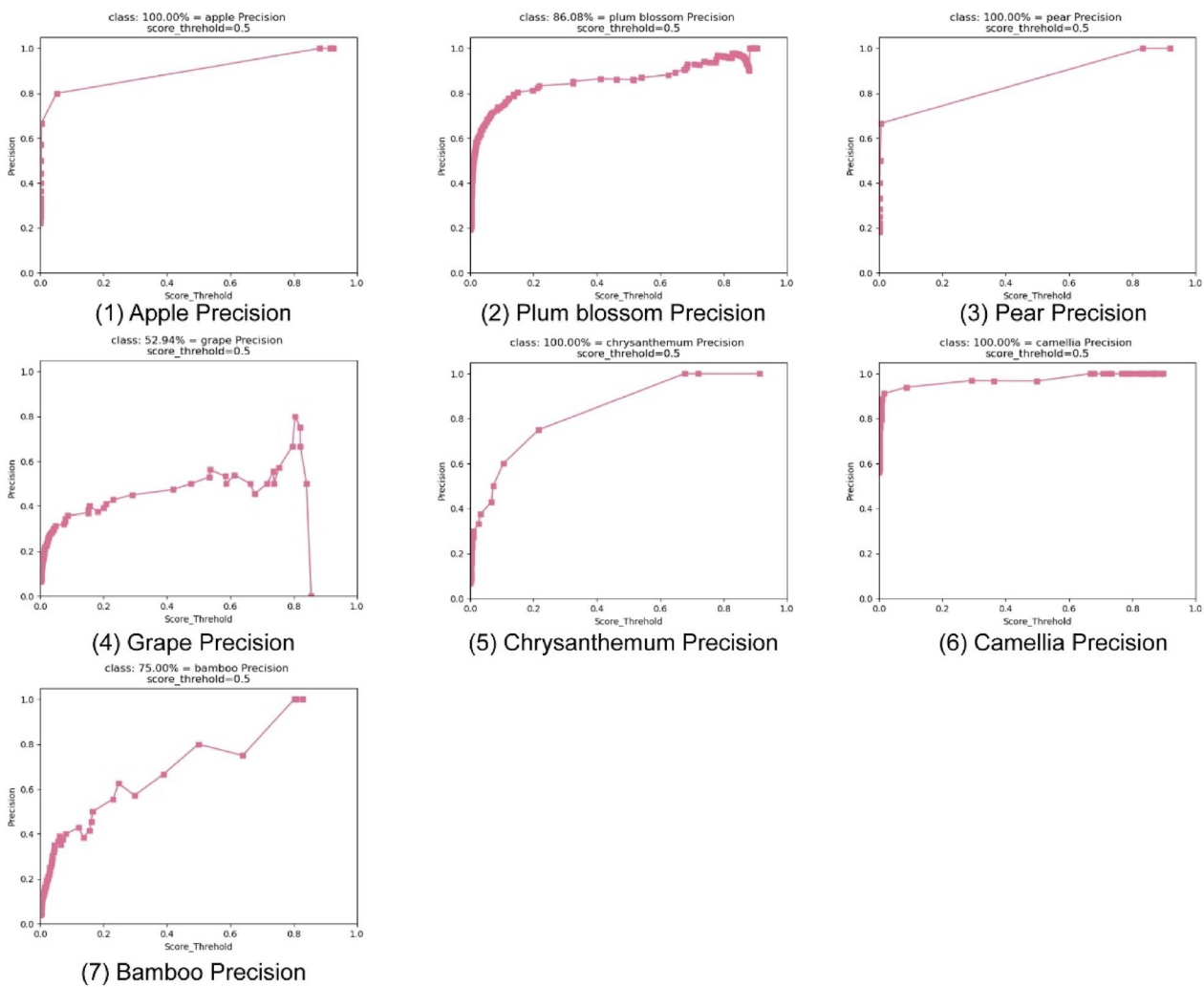
**Fig. 11** Model F1 score statistical analysis

indicates that the model produces more false positives when predicting bamboo.

- (4) Recall is a key indicator for evaluating the detection capabilities of a model. It measures the ratio of positive category samples identified by the model to the actual positive category samples. The following trend can be seen from Fig. 13: the recall rate of apples, chrysanthemums, and pears reaches 100% when the score threshold is 0.5. This means that the model is able to detect all true instances of the apple and pear categories without missing a detection. The recall rate of plum blossoms is close to 96%, which is a high recall rate, which means that the model also performs well in this category, and only a very small number of real plum blossoms are missed. The recall rate of grapes is low and fluctuates greatly, which indicates that the model's

detection ability for the grape category is not stable enough. Although the recall rate of camellia does not reach 100%, it remains at a high level (about 90%), which shows that the model has good detection ability for this category. Bamboo has the lowest recall rate, and the recall rate drops sharply as the score threshold increases. This shows that the model misses many instances when detecting the bamboo category.

- (5) The mean average precision (mAP) and the log-average miss rate are shown in Fig. 14. mAP is a comprehensive indicator of model detection performance that calculates the average of the average precision (AP) across all categories. In this model, the mAP is 85.97%, which is quite a high score and indicates that the model has good detection capabilities in most categories. As can be seen from



**Fig. 12** Precision statistical analysis

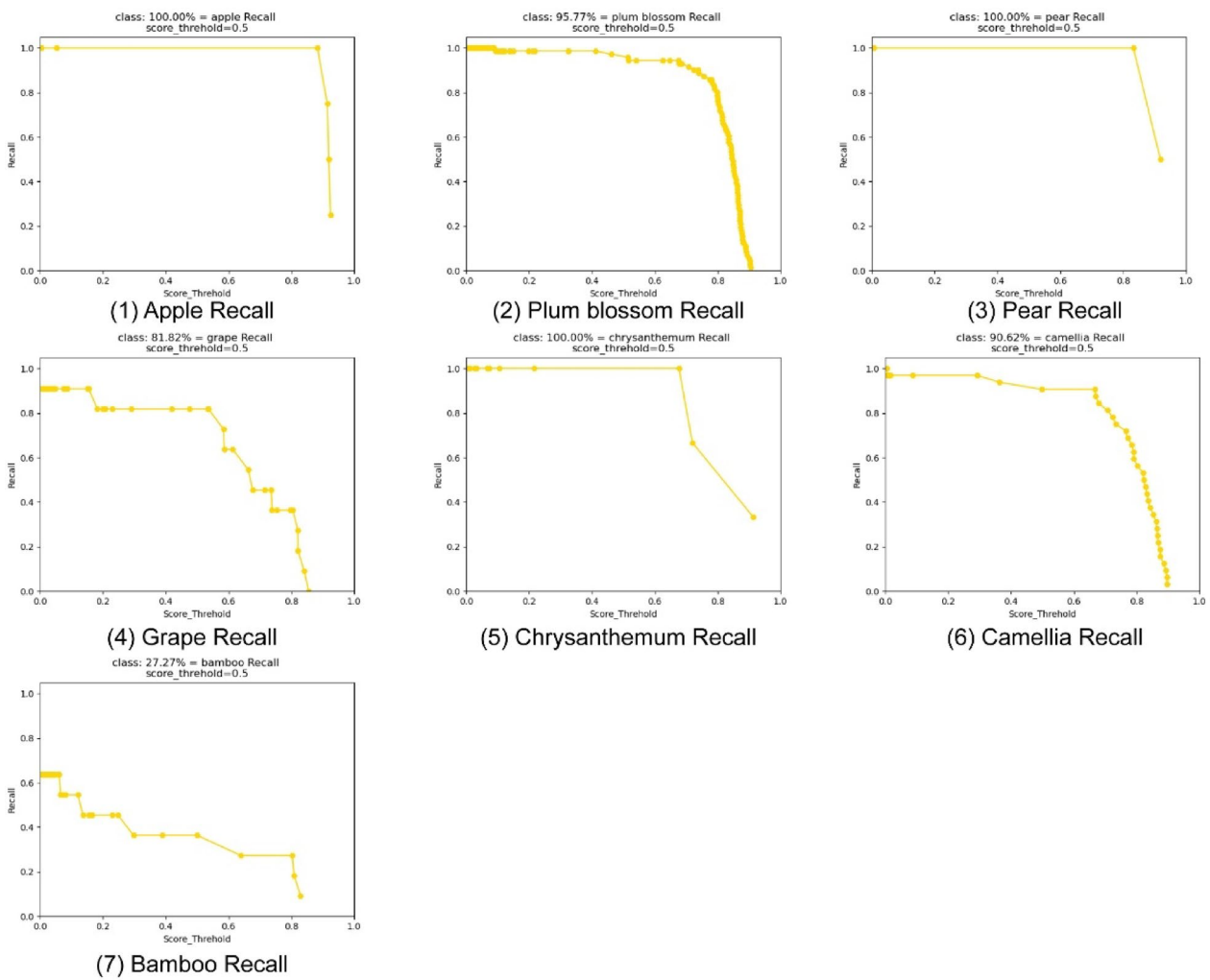
Fig. 14, except for the relatively low AP of grapes (0.58) and bamboo (0.48), the APs of other categories such as pears, chrysanthemums, apples, camelias, and plum blossoms are close to or reach perfect 1.0. The log-average miss rate is the logarithmic average of the miss rates measured at different false-positive rates. A lower log-average miss rate means better detection performance. Grapes and bamboo have higher miss rates, 0.71 and 0.68, respectively, while plum blossoms have a smaller but non-zero miss rate (0.16). The miss detection rate for other categories is zero, indicating that the model detects almost no false negatives for these categories.

and consistent characteristics of these two categories. Although plum blossoms perform well in precision and recall, there is still room for slight improvement in the log-average miss rate, indicating that there are still a small number of plum blossoms missed by the model. Grapes and bamboo perform poorly in various evaluation indicators, especially in log-average miss rate, which may reflect the high variability of samples in these categories. Overall, the model's performance highlights the importance of employing a diverse and broad range of training samples during training, which helps the model better generalize and identify various Chinese porcelain inlay categories.

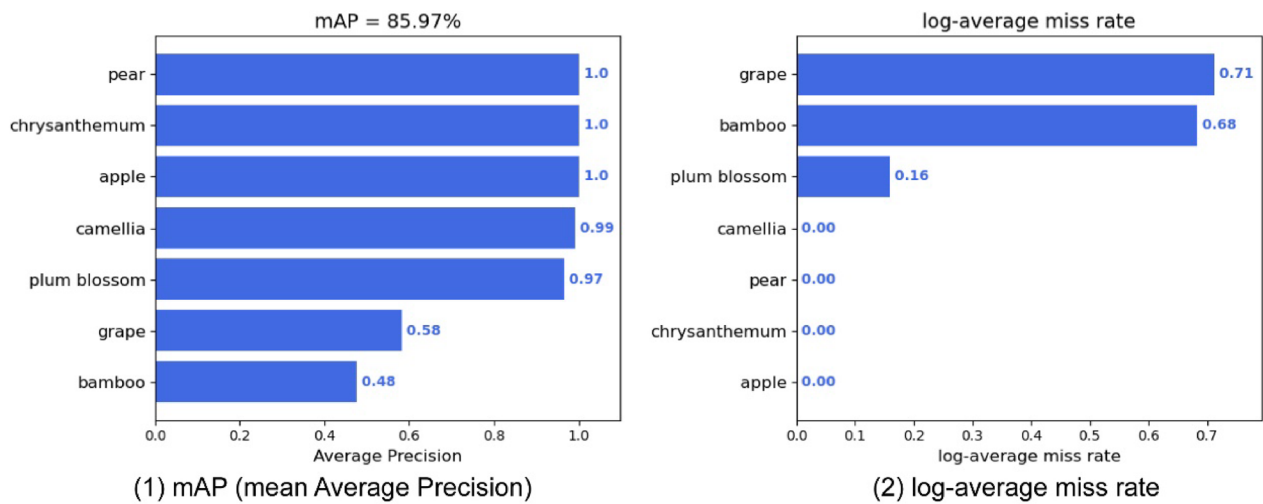
**Model testing**

To sum up, the apple and pear categories perform well in all indicators, which is related to the obvious

In the model test, the recognition results of the YOLOv4 model used in this study are analyzed at each epoch. As



**Fig. 13** Recall statistical analysis



**Fig. 14** The mean average precision (mAP) and the log-average miss rate statistical analysis

shown in Fig. 15, the test set contains multiple categories of Lingnan Chinese porcelain inlay patterns, including plum blossoms, pears, grapes, chrysanthemums, camellias, bamboos, and apples. The recognition results of each category correspond to three key epochs in the training process: the maximum training period (Max Epoch), the minimum training loss (Min Loss), and the minimum verification loss (Min Val Loss). To evaluate the recognition performance of the model, different confidence thresholds are also used to filter the recognition results, including confidence settings of 0.3 and 0.1.

The test results show that (1) the model has the best overall recognition effect in Min\_Val\_loss (the 116th epoch). This is consistent with the lowest point of verification loss observed in the previous training stage. This is significantly manifested in the recognition of bamboo-type Chinese porcelain inlays. The model with the minimum verification loss point can identify increasingly accurate targets. (2) Under the 0.3 confidence threshold, all models have high recognition accuracy for categories such as plum blossoms, pears, grapes, and chrysanthemums, and the bounding boxes delineate the targets compactly and accurately. This shows that the model has better generalization ability and localization accuracy in these categories. (3) At the maximum epoch (the 200th epoch), although the training loss reaches a low level, the model's recognition performance in some categories showed a slight decline. Bounding boxes such as bamboo and apples exhibit some misrecognition, which may be due to model overfitting on these categories.

To further test the accuracy of model recognition, the above test results are compared with those of the better Min\_Val\_loss model and tested under a more stringent 0.1 confidence threshold. The results show that for the plum blossom, pear, grape, and chrysanthemum pattern types, the model is still highly accurate. The number of missed bamboo inspections decreases. However, some correct targets are misidentified, such as the apple-patterned type of Chinese porcelain inlay. This shows that lower confidence thresholds can reduce missed detections but also lead to false detections of some Chinese porcelain inlay pattern types. In practical applications, the model needs to find a balance between accuracy and misidentification.

In addition, for decorative and complex categories such as camellia, the model shows better recognition ability at the 116th epoch (showing that a tiny camellia was also recognized in the bamboo picture). This may be due to the generalization performance of the model during this training stage. In practical application of the model, this study can adjust the confidence threshold according to the characteristics of different categories to obtain the best recognition effect. From the

above-detailed analysis of the test results, it can be seen that the model of the 116th epoch reaches a balance in both loss value and detection performance. This model is used for the next in-depth analysis.

To further test the model, the detection head output of the model in the 116th epoch is used to generate a heatmap (Fig. 16). These heatmaps show the areas where the model focuses its attention when processing test images. By parsing these heatmaps, researchers can gain deep insights into how models process and identify different features. In the heatmap in Fig. 11, layer 0 shows that the model initially focuses on high-contrast areas in the image, which are typically associated with the body and boundaries of the inlay pattern. For example, in the "Bamboo" category, the hotspots are mainly concentrated on the tips and intersections of bamboo leaves, which are the most distinguishing visual features of bamboo leaves. In layer 1, the model's focus begins to shift to more detailed features, such as subtle changes in texture and pattern, which means that the model analyzes the internal structure of the Chinese porcelain inlay pattern more finely. In layer 2, the model's area of concern is further expanded to encompass the entire pattern. This may be a comprehensive evaluation process before the model performs category classification.

Score heatmaps in the detection head output provide intuitive information about the probability that the model predicts the presence of a target in a specific area. According to the class output of head 0, the model predicts probability distributions for different classes based on different regions of the image. For example, for the chrysanthemum category, in the category output of head 1, hotspots are densely distributed at the edges of chrysanthemum flowers, indicating that the features of these areas are crucial for category judgment.

When the category score (class\_score) is output, the hot spots are not only concentrated on the target object but also generate responses in some nontarget areas, suggesting that the model has potential misidentification risks in these areas. For example, in the output of the category score of head 2, some background areas also show slight heat, possibly due to some visual similarities between the background and the Chinese porcelain inlay pattern. In summary, through this in-depth heatmap analysis, this study confirms the model's recognition capabilities for specific categories and reveals how the model processes visual information in complex scenes.

To test the effect of the model in practical applications, the research team took a picture of a new Chinese porcelain inlaywork from the scene. A section relevant to this study was selected to demonstrate the model's performance on images it has never seen. Figure 12 shows

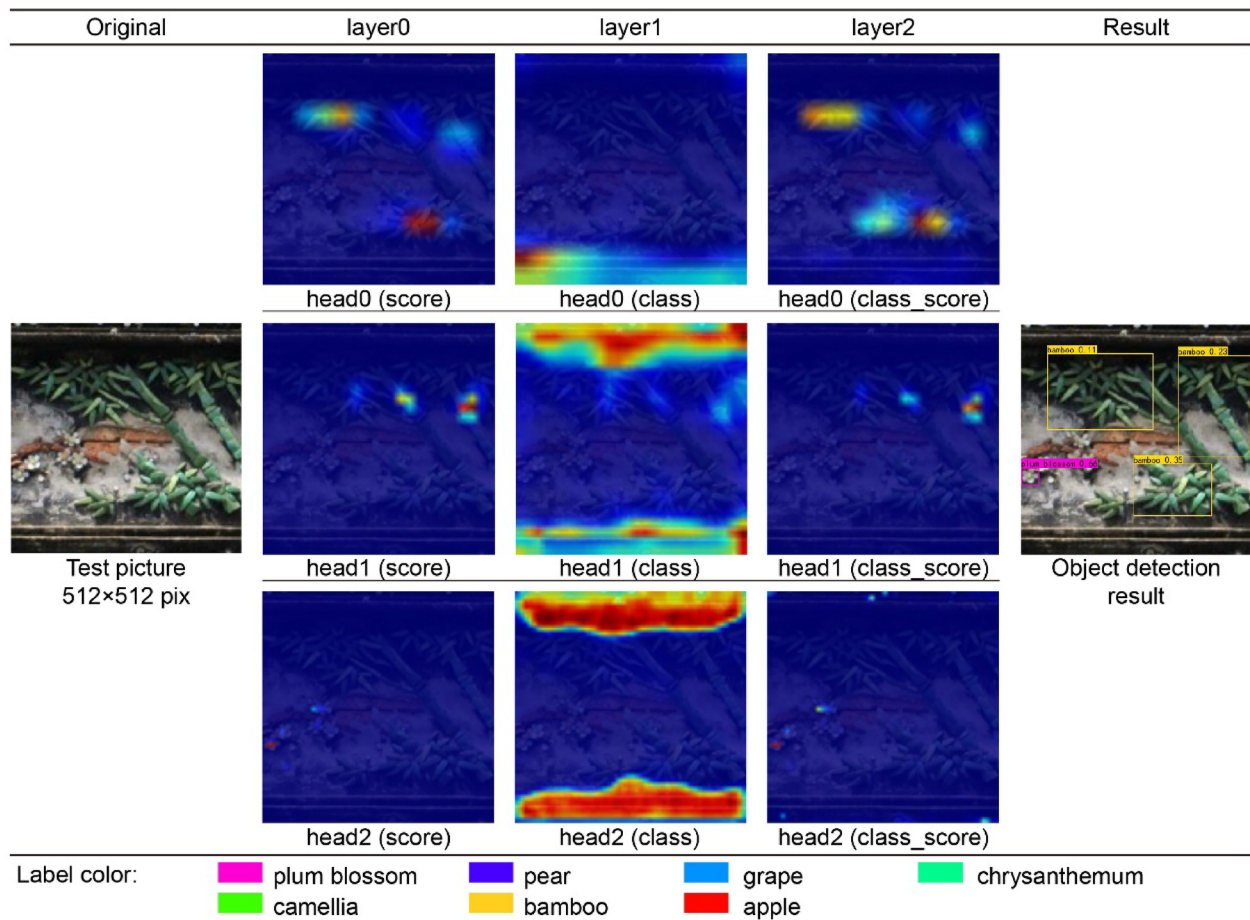
Category	Original	Max Epoch	Min Loss	Min Val loss	
		0.3 confidence	0.3 confidence	0.3 confidence	0.1 confidence
plum blossom					
pear					
grape					
chrysanthemum					
camellia					
bamboo					
apple					

Label color: ■ plum blossom    ■ pear    ■ grape    ■ chrysanthemum  
■ camellia    ■ bamboo    ■ apple

**Fig. 15** Results of model testing on different types of Chinese porcelain inlay images

the application results, detailing the response of the model at different levels and detection heads, as well as the final object detection results. Figure 17 shows the complex Chinese porcelain inlay pattern, which includes

multiple categories involved in the study, such as plum blossoms and camellias. Additionally, the size is not the standard 512×512 pixels but a banner image captured directly by the camera. The model's response in layer 0



**Fig. 16** Heatmap results of the detection head during the Chinese porcelain inlay pattern model detection process

shows an initial response to high-contrast features in the image. These highlighted areas indicate potential target locations. In the score output of head 0, the model shows high attention to the edges of the Chinese porcelain inlay part, which shows that the model can initially identify the dividing line between the pattern and the background.

As the level increases, the focus of the model becomes more detailed. In layer 1, the class output of head1 shows that the model begins to distinguish different Chinese porcelain inlay patterns. The heatmap shows obvious bright spots on specific patterns, indicating the sensitivity of the model to different pattern features. In layer 2, the class\_score output of head 2 shows that the model's attention is focused on specific patterns, such as plum blossoms and camellias, and these areas appear as unique color patterns on the heatmap, clearly distinguishable from other patterns.

In the object detection results, the model successfully marks the location of the target Chinese porcelain inlay patterns in the original image and identifies different categories with bounding boxes of different colors. These

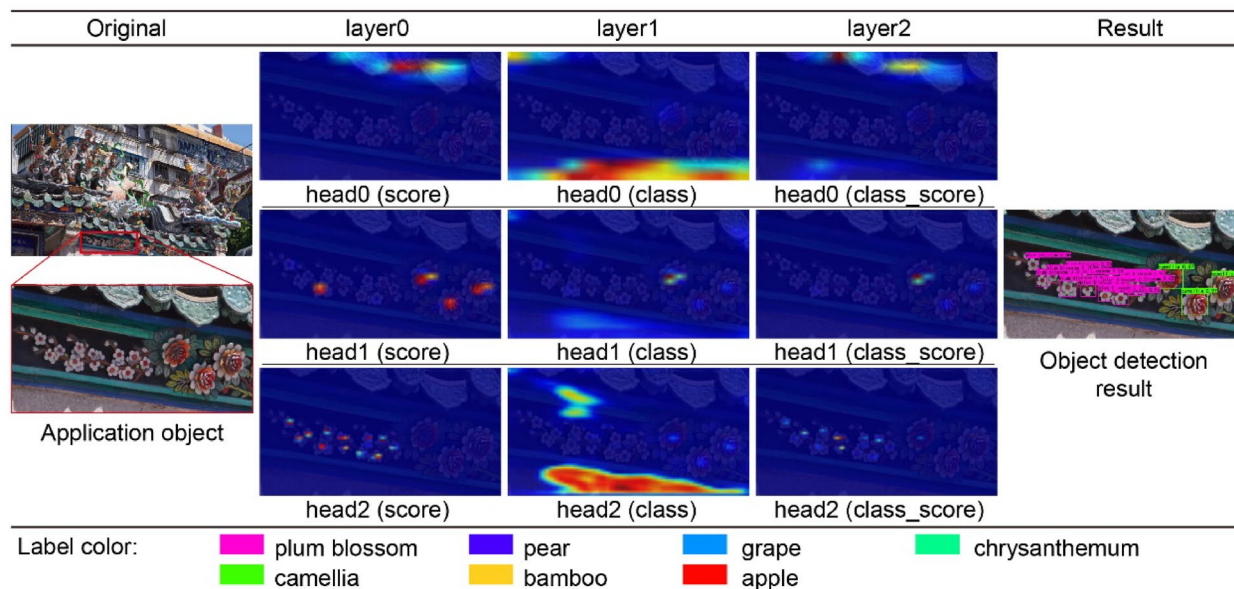
detection results correspond to the heatmap analysis, verifying the effectiveness of the model in actual application scenarios. The model can accurately identify and locate Chinese porcelain inlay patterns from complex backgrounds, demonstrating its potential for practical applications. Through the above tests, the model's application ability in real scenarios is verified and adaptability to the Lingnan Chinese porcelain inlay pattern identification task is demonstrated. This result not only confirms the practicality of the model but also provides future application prospects for on-site cultural heritage protection and digitization.

**Discussion: application and accuracy statistics of the model in traditional architectural scenarios**

**Model application in traditional architectural scenarios**

The ultimate goal of researchers is to accurately identify the name of each style from an image containing a large number of porcelain inlay styles. Therefore, researchers need to apply model testing to verify the accuracy of the experiment. The detection results of the test set are





**Fig. 17** Model detection and test results for new Chinese porcelain inlay images

shown in Fig. 18. The researchers rephotographed some photos from traditional Lingnan architectural decorations as materials for model application; these materials ranged in size. In the actual scene application of traditional architecture, the research revealed that (1) the model can more accurately detect Chinese porcelain inlay image styles in different positions, different lights, and different modeling techniques. (2) This model is compatible with Chinese porcelain inlay images of different sizes, angles, and definitions. (3) This model can eliminate interference from other elements in the photo and can still accurately identify the Chinese porcelain inlay plant style when encountering mixed subjects such as animals and people. It can be concluded that the YOLOv4 algorithm has good adaptability for detecting Chinese porcelain inlay images.

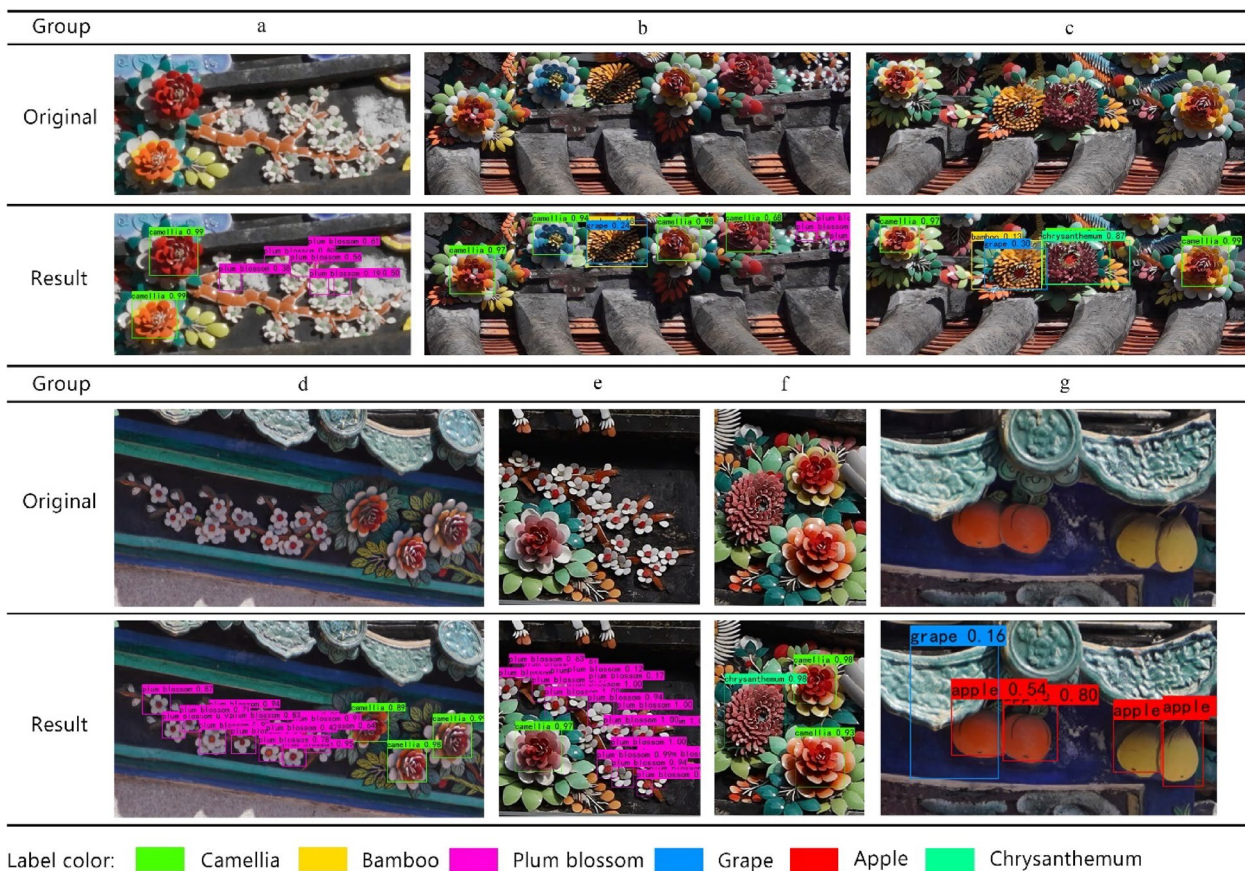
Figure 18 shows the test photos and test results of the Chinese porcelain inlay, a traditional Lingnan architectural decoration. We can see that the overall detection effect of this model is acceptable. Most Chinese porcelain inlay-style images can be accurately identified, especially for camellias and plum blossoms, which can achieve high detection accuracy. However, images of bamboo, grapes, and other styles with small sample sizes are confused, especially when the image is dark. This can be seen in groups b, e, and g. In summary, the YOLOv4 algorithm in this study can effectively detect the Chinese porcelain inlay style of traditional Lingnan architectural decoration, but the detection accuracy needs to be improved and cannot completely replace human visual inspection. However, this model can improve people’s understanding

of the intangible cultural heritage of Chinese porcelain inlay to a certain extent. Understanding culture is particularly important now that this craft is declining. Similarly, there are numerous other intangible cultural heritages and architectural decorations similar to Chinese porcelain inlays. In the future, this research can also be applied to protecting cultural heritage and modern architecture and design in other countries and regions.

**Comparative analysis with YOLOv8**

To evaluate the performance of YOLOv4 and its similarities and differences with YOLOv8, the researchers calculated a total of 581 successfully tested and missed labels by these two models. To more clearly understand the model’s detection capabilities versus its classification performance, the researchers calculated the accuracy and error rate against the total number of test labels, which represent the labels where the model made a classification, while the missing rate was calculated against the combined count of test labels and missed labels to assess the proportion of labels that the model failed to detect out of the potential detections. According to the data in Table 3, researchers can draw the following conclusions:

- (1) When the Chinese porcelain inlay image is relatively regular, the YOLOv4 model is better. YOLOv4 classifies chrysanthemum, plum blossom, camellia and apple with accuracies of 87.5, 70.99, 80.39, and 80%, respectively, while YOLOv8 correctly classifies these categories at 58.82, 65.38, 70.83, and 37.5%,



**Fig. 18** The detection effect of different Chinese porcelain inlay project site photos

respectively. This is much lower than the results of the YOLOv4 model. The possible reason is that the images of chrysanthemums, plum blossoms, camellias, and apples are relatively regular, with round outlines and similar and layered internal petals. Therefore, under the condition that Chinese porcelain inlay images are relatively regular, YOLOv4 has better recognition capabilities than YOLOv8.

- (2) When the Chinese porcelain inlay image is irregular, the YOLOv8 model is better. YOLOv8's classification accuracy for bamboo, grapes, and pears is 75.00, 62.50, and 80.00%, respectively, which is much higher than YOLOv4's 12.50, 33.33, and 50.00%. The possible reason is that the images of bamboo, grapes, and pears are mostly shaped, and the overall outline is irregular. Therefore, YOLOv8 has better classification ability than YOLOv4 under the condition of irregular Chinese porcelain inlay images.
- (3) When the Chinese porcelain inlay image is rich in color, the YOLOv4 model is better. In the process of making chrysanthemums, plum blossoms, camellias, and other inlay porcelain flowers, Chi-

nese porcelain inlay craftsmen often use not just one color but multiple colors layered on top of each other to make the work more gorgeous. YOLOv4 has a lower error rate than YOLOv8 in classifying these gorgeous flowers. Therefore, under the condition that Chinese porcelain inlay images are rich in color, YOLOv4 has better classification capabilities than YOLOv8.

- (4) In poor light conditions, the YOLOv8 model is better. YOLOv8 outperforms YOLOv4 in terms of data loss rate in most categories, especially apples and pears. The possible reason is that apples and pears tend to be decorated in more remote parts with poor lighting, so YOLOv8 is predicted to have a stronger ability to detect images with poor lighting.
- (5) When Chinese porcelain inlay images overlap a lot, the detection capabilities of the YOLOv4 and YOLOv8 models are similar. In the detection of plum blossoms, the indicators of the YOLOv4 and YOLOv8 models are relatively close. The possible reason is that plum blossoms often appear in groups, and their petals are often stacked, causing

**Table 3** Manual Validation Comparison between YOLOv4 and YOLOv8 in Chinese porcelain inlay

	Model	Chrysanthemum	Plum blossom	Camellia	Bamboo	Grape	Apple	Pear
Label quantity	YOLOv4	8	131	51	24	8	5	2
	YOLOv8	19	107	27	4	24	9	5
Accuracy quantity	YOLOv4	7	93	41	3	3	4	1
	YOLOv8	10	93	17	3	15	3	4
Accuracy rate	YOLOv4	87.50%	70.99%	80.39%	12.50%	33.33%	80.00%	50.00%
	YOLOv8	58.82%	65.38%	70.83%	75.00%	62.50%	37.50%	80.00%
Quantity of errors	YOLOv4	1	38	1	21	6	1	1
	YOLOv8	9	14	10	1	9	6	1
Error rate	YOLOv4	12.50%	29.01%	1.96%	87.50%	66.67%	20.00%	50.00%
	YOLOv8	41.18%	34.62%	29.17%	25.00%	37.50%	62.50%	20.00%
Missing quantity	YOLOv4	5	57	27	4	1	7	4
	YOLOv8	1	43	2	1	2	1	1
Missing rate	YOLOv4	38.46%	30.32%	34.62%	14.29%	10.00%	58.33%	66.67%
	YOLOv8	5.56%	16.13%	7.69%	20.00%	7.69%	11.11%	16.67%

Source: author statistics

machine learning to mistakenly regard multiple plum blossoms as one.

- (6) When the Chinese porcelain inlay image element is lost, the detection capabilities of both the YOLOv4 and v8 models are insufficient. In the detection of plum blossom petals and grapes, due to their small size and limited quantity, especially in the neglected Chinese porcelain inlay works, these elements tend to fall off, which brings challenges to the detection. Once displacement occurs, it can be difficult to detect it. Therefore, both models are not good at detecting stacked or missing elements in the test data.

In general, YOLOv4 and YOLOv8 have their own advantages and disadvantages in Chinese porcelain inlay image classification, and it is impossible to choose based on the advanced version alone. However, researchers learned from long-term field surveys that the survival of the Chinese porcelain inlay, an intangible cultural heritage, is very difficult. Today, as urbanization accelerates, there are many high-rise buildings, a large number of ancestral halls and temples have been banned, and the carrier on which Chinese porcelain inlay relies for its survival has gradually disappeared. In addition, the economic source of Chinese porcelain inlay craftsmen is very unstable, and young people are unwilling to engage in this industry because of the high time cost and low economic return. According to Mr. Xu Shaopeng, the inheritor of Chinese porcelain inlay, not only is it difficult to recruit Chinese porcelain inlay apprentices, but they also have to be paid daily, which makes life even worse

for Chinese porcelain inlay craftsmen who already have a difficult life. According to Mr. Lu Boxin, the inheritor of Chinese porcelain inlay, making Chinese porcelain inlay is purely out of emotion and love and does not make any money. In order to promote Chinese porcelain inlay skills, they have to pay high stall fees to participate in some exhibitions. Therefore, they mainly support their main business of Chinese porcelain inlay by doing some side jobs on construction projects. Relevant cultural relic protection departments and university scholars in China attach great importance to the protection of Chinese porcelain inlay, an intangible cultural heritage. However, due to the political system, these departments are not qualified to generate income and can only rely on financial allocations to maintain basic operations. Especially in the context of the poor overall economic environment in the world in recent years, it is not able to provide excessive financial, technical, and talent support to Chinese porcelain inlay. Therefore, in the actual operation stage, we should consider the cost of technology and other issues. The training of YOLOv8 is relatively slow. It requires a 3G environment package when calling the model, requires a higher-configuration training environment, and consumes a lot of computer resources. YOLOv4 can be called directly without an environment. The application is not much different from YOLOv8. Compared with YOLOv8, YOLOv4 has a heat map, which can clearly show the effectiveness of our presentation mechanism. In the comprehensive comparative analysis in this section, YOLOv4 can already meet the basic requirements for Chinese porcelain inlay image recognition, and there is no need to invest more in upgrading the equipment. In addition, YOLOv4 has been operating for a long time, the

model is stable, and it is relatively mature in all aspects. However, the new model will have many uncertainties. Therefore, YOLOv4 is the preferred model for the specific dataset and category being analyzed. However, the performance of each version may vary depending on specific use cases and dataset characteristics, and there is still considerable room for improvement, requiring extensive training to improve its accuracy.

## Conclusion

### Research discovery

With rapid development, the development space of traditional architecture has become increasingly narrow, and the inheritance of Chinese porcelain inlay craftsmanship has gradually declined [38]. Chinese porcelain inlay is currently facing the risk and pressure of inheriting “people, technology, and objects” across generations. Increasing examples of cultural heritage similar to Chinese porcelain inlays are on the verge of being lost. If people cannot understand Chinese porcelain inlay culture and correctly understand intangible cultural heritage, this will be a cruel reality that cannot be ignored in Lingnan culture. Therefore, it is necessary to use digital means to protect Chinese porcelain inlay, a national intangible cultural heritage. The complicated patterns in traditional architectural decorations in the Lingnan area make identification difficult, and superficial sightseeing tours cause people to lack a deep understanding of the culture.

Therefore, this study optimized the Lingnan Chinese porcelain inlay image recognition model based on YOLOv4 through innovative techniques. The model used a thousand first- and second-hand images collected from field surveys and the internet for preprocessing, style classification, and annotation and then conducted rigorous model training. After conducting model testing to verify the accuracy of the experiment, we found that the YOLOv4 algorithm has good adaptability in the Chinese porcelain inlay pattern detection task and achieves good results in practical applications. Compared with manual identification, this method is timelier and more universal and plays an important role in people’s understanding of Chinese porcelain inlay culture. The results were mainly obtained in the following three aspects:

- (1) Efficient identification of Chinese porcelain inlay types: the model in this study was specifically adjusted to effectively identify multiple Chinese porcelain inlay types, including traditional Lingnan inlay porcelain patterns such as plum blossoms and camellias. In the 116th epoch, the model showed excellent generalization ability, and the verification loss reached the lowest value of 0.88. The lowest training loss in the 195th epoch was 0.99, indicating

that the model reached an optimal balance point for both recognition accuracy and processing speed.

- (2) Fine-grained visual feature extraction: the heatmap analysis applied in this study reveals the model’s ability to extract fine-grained visual features at different levels. Through careful interpretation of the detection head output, the study revealed that the model can distinguish subtle details in the Chinese porcelain inlay pattern, such as texture and color differences, which is difficult to achieve with traditional methods.
- (3) Application verification in actual scenarios: the model proved robust in application tests for new materials. On randomly selected on-site Chinese porcelain inlay images, the model not only successfully identified the target category but also accurately located the pattern boundary. This result verifies the model’s ability to handle complex visual scenes in real-life environments and demonstrates its application potential in the field of digital protection of cultural heritage.

In summary, the innovation of this study is to specifically adjust and optimize the YOLOv4 model to adapt to the complexity of the Lingnan Chinese porcelain inlay pattern, achieving significant improvements in accuracy, speed, and generalizability. Future work will continue to explore further enhancing the model’s application capabilities in a wider range of cultural heritage image recognition tasks.

### Limitations and future work

This article is based on the machine learning model of YOLOv4 architecture, which plays a decisive role in the real-time detection and identification of the traditional Chinese porcelain inlay style in the Lingnan area and the protection of cultural heritage. However, some limitations remain, which may make the identification efficiency uncertain. Purely handmade architectural decorations, such as Chinese porcelain inlays, have a strong personal touch. Even with the same pattern and style, different craftsmen have different production methods and habits. Some “freehand-style” craftsmen refer to the painting style of Chinese paintings when creating, which is very different from the traditional style. This leads to difficulties in the recognition of machine learning models and cannot include all cases. In addition, the quality of the training samples has a greater impact on the training of the YOLOv4 model. Poor lighting conditions and low-definition image quality will affect the accuracy of the model. Therefore, subsequent research also needs to collect a large number of clear images for machine learning model training. In future work, researchers can also

use drones, telephoto lenses, and other equipment to record Chinese porcelain inlay images at closer distances and from multiple angles, thereby improving the quality of the dataset.

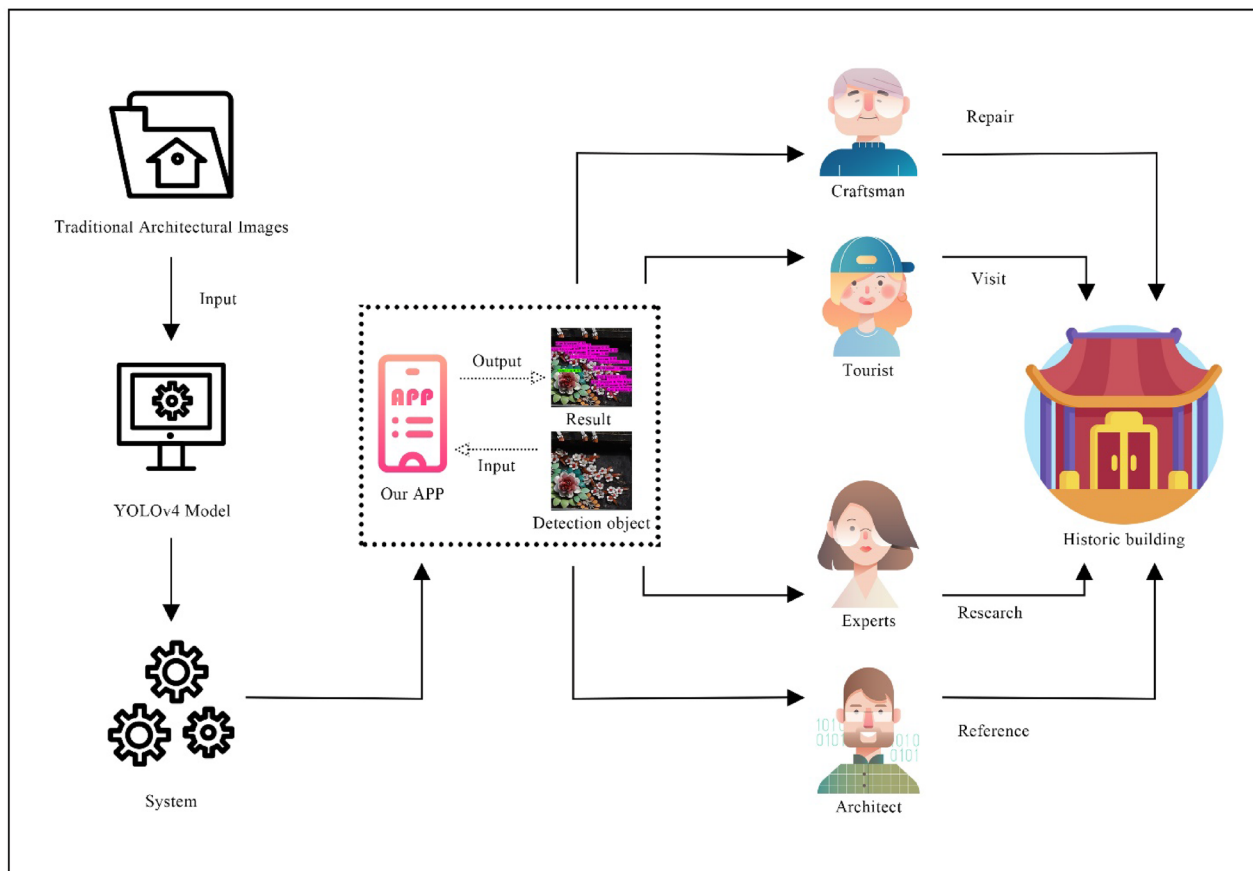
Based on the deficiencies and limitations identified in our study model, future research can attempt improvements in the following areas:

- (1) Enhancement of feature extraction capability: The integration of more advanced feature extraction modules, such as attention mechanisms (e.g., CBAM or SE modules), into the CSPDarkNet53 backbone network can enhance the model's ability to recognize important features. This will help the model focus better on key parts of the image, thereby improving the understanding of complex scenes and accuracy in object detection.
- (2) Optimization of the feature fusion strategy: Although YOLOv4 already employs the PANet and SPP structures for feature fusion, there is still room for improvement. Exploring new feature fusion techniques, such as improved versions of the fea-

ture pyramid network (FPN) or using more efficient methods of upsampling and downsampling, can further enhance the efficiency and effectiveness of feature fusion across different scales.

- (3) Research and customization of loss functions: The existing loss functions may have limitations for specific tasks. By researching and developing loss functions with stronger specificity, such as improved IoU loss or loss functions that incorporate target size and class imbalance, the model's accuracy in object localization and classification could be further improved.

Researchers are also preparing to extend the application of this technology to cultural heritage protection work in different regions and fields and provide a new perspective for modern architecture and design research. For example, researchers can develop an app for automatic image recognition of Chinese porcelain inlay styles based on the YOLOv4 model (Fig. 19). This app can have a certain impact on craftsmen, tourists, experts and scholars, architects, and designers.



**Fig. 19** Application for automatic Chinese porcelain inlay image recognition

- (1) For craftsmen, the application based on this model can quickly locate the damage style by identifying the patterns of Chinese porcelain inlays and contacting Chinese porcelain inlay craftsmen in a timely manner to provide repair materials or replacement parts for the pattern. The use of this tool will simplify the maintenance and repair process, speed up the assessment of damage, and allow faster restoration of ancient buildings and cultural relics.
- (2) Tourists who are coming into contact with Chinese porcelain inlays for the first time or who are lovers of Chinese porcelain inlay art can take pictures of specific Chinese porcelain inlay artworks or ancient buildings and identify the patterns in the pictures by using this app. The identified patterns will be linked to popular science introductions about the cultural connotations of Chinese porcelain inlays, such as its meaning, origin, production characteristics, and cultural implications, so people can learn new knowledge and gain a deeper understanding of Lingnan culture from them.
- (3) Scholars and experts can use this app to identify rare patterns, which will provide a starting point for further study of traditional Lingnan architectural decoration. The app can also show documents and pictures related to image recognition results so experts can deeply explore the connotations of cultural heritage, which will help them formulate cultural heritage protection strategies.
- (4) For architects and designers, when conducting on-site research on ancient buildings, using this app to identify the decorative patterns on the site can help them instantly understand their relevant cultural context. They can store these identification results in the app at the same time, so they provide research sources and inspiration for them in the design of ancient buildings and revival-style historical districts to maintain cultural authenticity in modern design.

In summary, despite the current technical constraints on the use of machine learning technology for identifying Chinese porcelain inlay patterns, there is promising potential for future improvements and applications. Future work will refine the accuracy of

the YOLOv4 model but will also explore new frontiers in the application of artificial intelligence technology for cultural heritage preservation, bridging the gap between traditional craftsmanship and modern technological capabilities.

### **Appendix A: the spread of Chinese porcelain inlay as architectural decoration overseas**


#### **See Table 4 and Fig. 20 here**

In addition to the Fujian and Chaoshan areas of Guangdong in China, there are also many traditional Chinese buildings overseas that use Chinese porcelain inlay for decoration. It is also widely distributed in Vietnam, Thailand, Myanmar, Japan and other countries. Chinese porcelain inlay not only appears in traditional religious temples but also in Chinese former residences, chambers of commerce, and even Chinese cemeteries and worship-type buildings. On the other hand, there are some manufacturers that specialize in mass-producing porcelain tiles of various specifications, shapes, and colors for Chinese porcelain inlay craftsmen, which greatly reduces the labor cost of Chinese porcelain inlay production and improves work efficiency. But the "craftsman spirit" has been lost to a certain extent. Nowadays, in order to reduce costs, craftsmen in Southeast Asia mostly purchase finished Chinese porcelain inlay for assembly. Only a small number of ancient buildings with unique commemorative significance still go all the way to China to request craftsmen to design and produce Chinese porcelain inlay locally. For example, the Chinese porcelain inlay inheritors of the Chaozhou Chinese Porcelain Inlay Museum are planning to restore the Chinese porcelain inlay works at the House of Tan Yeok Nee in Singapore, under strong invitation from the local government.





### **Appendix B: machine learning runtime environment**

Machine learning environment: the operating system is Windows 11 (X64), the CUDA version is 11.5, the deep learning framework is PyTorch (1.13.0), and the graphics card and processor are a GeForce GTX 3070 (16 G) and an AMD Ryzen 9 5900HX (3.30 GHz), respectively.

**Table 4** Incomplete statistics of overseas buildings using Chinese porcelain inlay as decoration

Country	Name (English)	Name (other language)	Photo	Location	Years	Historical characteristics
Thailand	Wat Bowonniwet Vihara	วัดบวรนิเวศวิหาร		Phra Sumen road, Boworn Niwet, Phra Nakhon	1824	one of Bangkok's most important temples
	Kuan An Keng Shrine	ศาลเจ้าแม่กวนอิมเก็นอันแกง		Wat Kanlaya, Thonburi, Bangkok 10600	1848	One of the oldest shrines in Thonburi
	Wat Mangkon Kamalawat	วัดมังกรกมลาวาส (เส่งนयी)		432 Charoenkrung Rd, Pomphab Sat-tru Fai, Bangkok	1871	an emblem of Chinese architecture and the religious heart of the Chinese population living in Bangkok
	Wat Cheen Pracha Samosorn	วัดจีนประชาสโมสร		291 Supakit road, Na Muang, Muang, Chachoengsao	1873	One of the most important temples in Chachoengsao province
	Pueng Thao Kong Shrine	ปู่แก่ก		64–66, Praisanee Road Chang Moi, Chiang Mai	1879	One of the most popular temples in Chiang Mai
	Sang tham Shrine	ศาลเจ้าตงก้องต๋อ		72/1 Yaowarat Rd, Amphoe Mueang Phuket, Chang Wat Phuket	1891	granted the Conservation award by princess Maha Chakri Sirindhorn

**Table 4** (continued)

Country	Name (English)	Name (other language)	Photo	Location	Years	Historical characteristics
	Wat Mangkon Buppharam	วัดมังกรมปไพโรสถาราม		Philo, Laem Sing District, Chanthaburi 22190	1977	have a history of Phra Ajarn Jean Wang Sawattthiwat (Sok Heng)
Malaysia	Cheng Hoon Teng Temple	青云亭		No. 25, Jalan Tokong, 75200 Melaka	1673	The oldest traditional Chinese temple in Malaysia
	Penang Kong Hock Keong Kwan Yim Teng	檳城广福宫观音亭		30, Jalan Masjid Kapitan Keling (Pitt Street), Penang	1800	The earliest place where Chinese Buddhism spread to the south in the 18th century; the oldest temple in Penang
	Leong San Tong Khoo Kongsi	龙山堂邱公司祠庙		18 Cannon Square, 10200 Penang	1850	The most unique Chinese clan association in Malaysia
	Penang Ban Ka Lan Fu Xing Gong Snake Temple	檳城万树兰福兴宫蛇庙		Jalan Tokong Ular, Jalan Sultan Azlan Shah, Bayan Lepas, Pulau Pinang	Before 1850	One of the top ten tourist attractions in Malaysia
	She Tek Tong Cheah Kongsi	世德堂谢公司祠堂		8, Armenian Street, 10200 George Town, Penang	1858	One of the oldest Fujian clan halls in George Town, Penang
	Penang Seng Hong Beow Temple	檳城城隍庙		Jalan C.Y.Choy, Penang	Before 1862	Chinese porcelain inlay is used as a beautiful place for architectural decoration



**Table 4** (continued)

Country	Name (English)	Name (other language)	Photo	Location	Years	Historical characteristics
	Hiang Tiang Siang Ti temple	玄天上帝庙		Carpenter Street	Before 1863	
	Hock Teik Cheng Sin Temple	福德正神庙		57, Lebuh Armenian, 10200 George Town, Penang	1892	One of the first Hokkien institutions established before the large influx of new immigrants to Penang
	Kuan Yin temple of Klang	巴生观音寺		Jalan Raya Barat	1892	Chinese porcelain inlay is used as a beautiful place for architectural decoration
	Cheong Fatt Tze	张弼士故居		the corner of Leith Street and adjacent to Muntri Street	1897	UNESCO Conservation and Heritage Award; the largest existing garden-style residence in Malaysia
	Kuala Lumpur Guan Yin Si	威镇宫观音寺		10, Jalan Maharajalela, 50150 Kuala Lumpur	1898	Chinese porcelain inlay is used as a beautiful place for architectural decoration
	Yap Kongsı and Ciji Temple	叶氏宗祠慈济宫		Lebuh Armenian, George Town, Penang Island	1920s	The heart of George Town, a UNESCO World Heritage Site

**Table 4** (continued)

Country	Name (English)	Name (other language)	Photo	Location	Years	Historical characteristics
	Kuala Lumpur Ju Xing Tang Guan Di Temple	吉隆坡怡保路聚星堂关帝庙		43-E, Jalan Mangga, Off Jalan Ipoh, Kuala Lumpur	1954	Chinese porcelain inlay is used as a beautiful place for architectural decoration
	Penang Batu Maung Cheng Chooi Keong	檳城峇都茅清水宮		Plot 9, Jalan Teluk Tempoyak, Batu Maung, Bayan Lepas, Pulau Pinang	1972	
	Kuala Lumpur Kepong Yuen Leong Temple	吉隆坡甲洞富城阮梁公圣佛庙		Lot 22942, Laman Rimbunan Kepong, Kuala Lumpur	1973	
	Thean Hou Temple	乐圣后天后宮		65, Persiaran Endah, Taman Persiaran Desa, 50460 Kuala Lumpur	1986	The largest Mazu temple in Malaysia
Singapore	Yueh Hai Ching Temple	粤海清庙		30B Phillip St, Singapore 048696	1826	2014 UNESCO Asia-Pacific Cultural Heritage Conservation Award Merit Award
	Thian Hock Keng	天福宮		158 Telok Ayer Street	1839	2001 UNESCO Asia-Pacific Heritage Cultural Heritage Conservation Award Honorable Mention
	Goh Chor Tua Pek Kong Temple	梧槽大伯公宮		249 Balestier Rd, Singapore 329708	1874	its use of red-painted plaster to simulate terracotta wall tiles

**Table 4** (continued)

Country	Name (English)	Name (other language)	Photo	Location	Years	Historical characteristics
	Tan Si Chong Su	陈氏宗祠		15 Magazine Road, Singapore	1876	Chinese porcelain inlay is used as a beautiful place for architectural decoration
	House of Tan Yeok Nee	陈旭年宅第		101 Penang Road	1882	
	Temple Cluster at Tampines Link and the former Hun Yeang Village	淡滨尼连接路和前洪阳村的寺庙群		Hun Yeang Village	1800s–1980s	
	Hoon San Temple	勸山寺		29 Jalan Lim Tai See	1902	
	Hougang Tou Mu Kung	后港斗母宫		779A Upper Serangoon Rd	1902	Listed as a national monument
	Lian Shan Shuang Lin Monastery	莲山双林禅寺		184 Jalan Toa Payoh	1904	Singapore's oldest Buddhist temple

**Table 4** (continued)

Country	Name (English)	Name (other language)	Photo	Location	Years	Historical characteristics
	Swee Kow Kuan	和升槽五王府爷清庙		46 Ang Mo Kio St. 61, S569161	1905	Chinese porcelain inlay is used as a beautiful place for architectural decoration
	Hong San See	凤山寺		31 Mohamed Sultan Road	1908	2010 UNESCO Asia-Pacific Cultural Heritage Conservation Award Excellence Award
	Leong San See Temple	龙山寺		371 Race Course Road Singapore	1917	Chinese porcelain inlay is used as a beautiful place for architectural decoration
	Shang Di Miao Chai Kong Temple	上帝庙		10 Marsiling Industrial Estate Road 3	1935	
	Hong Tho Bilw Temple	凤图庙		31 Marsiling Industrial Estate Road 3	1940s	
	Chu Sheng Temple	楚圣寺		48 Ang Mo Kio St. 61	1981	

**Table 4** (continued)

Country	Name (English)	Name (other language)	Photo	Location	Years	Historical characteristics
	Liuxun Sanhemiao	六巡三合庙		6 Ang Mo Kio St 63	1989	
	Tampines chinese temple	淡滨尼华人庙		25 Tampines street 21	1992	

Source: author statistics



**Fig. 20** Semi-finished product of Chinese porcelain inlay

**Appendix C: loss metrics during training**

In this study, the specific values of loss metrics during training are as follows:

Epoch	Training Loss	Validation Loss	Epoch	Training Loss	Validation Loss
1.0	172.13930274490127	12.873762130737305	101.0	1.1685655210587316	0.9897236009438832
2.0	10.460278093366695	6.93437074025472	102.0	1.1005525474429578	1.0584132785598437
3.0	7.576965894914212	5.6589510599772135	103.0	1.0754183641725914	1.1092751294374465
4.0	6.094109892845154	4.305464601516723	104.0	1.0620294294523118	0.9968639954924583
5.0	5.573051659684432	4.822180279095968	105.0	1.111012595376574	1.0450699046254157
6.0	5.133199482483971	4.105766387780507	106.0	1.11456208841357	0.9245148574312528
7.0	4.842857847088261	3.4729835669199627	107.0	1.1540739330694192	0.9442599579691887
8.0	4.7220738588419175	3.542416242758433	108.0	1.095662968219223	0.9130323206384977
9.0	4.312683376602661	3.454599134127299	109.0	1.0477591423611892	0.9282985327144464
10.0	4.174051762075353	3.868121687571208	110.0	1.135932919954447	0.9298349137107531
11.0	3.316101575592407	2.2222502768039702	111.0	1.1255582049769117	1.0815939399103323
12.0	2.916981518940818	2.2715669572353363	112.0	1.0196703025945146	1.0018715851008893
13.0	2.8403765158097545	1.9731369192401569	113.0	1.0689679367425746	1.0774113096296787
14.0	2.5171517320817576	2.0051982710758844	114.0	1.098165755937422	1.04351597627004
15.0	2.543552170794709	1.8176654641826948	115.0	1.1771025889013942	1.0278866517047087
16.0	2.4460136506118273	1.7437353074550628	116.0	1.0639248727295632	0.8759690086046855
17.0	2.336008021145835	1.7560623561342557	117.0	1.097059438719337	0.9746258775393168
18.0	2.172639416124588	1.5583937515815098	118.0	1.078225921811466	0.9685797038177649
19.0	2.3323311630384365	1.6672104666630427	119.0	1.0820308817844642	1.0399634127815565
20.0	2.2970562147018603	1.540772590537866	120.0	1.1518516897931135	1.0307459284861882
21.0	2.0715547736201967	1.596700624624888	121.0	1.1072028731661183	0.9630089787145456
22.0	2.075017137186868	1.5405688260992367	122.0	1.0817654501450689	0.9492419009407361
23.0	2.014883841114833	1.4985232830047608	123.0	1.1488866618365274	0.9928064301609993
24.0	1.8430926501190752	1.5739924470583597	124.0	1.0926379479916024	0.9000832661986351
25.0	1.741072788162339	1.5145430068174999	125.0	1.1091402412245148	1.0053825611869494
26.0	1.9046972470175951	1.3678817962606749	126.0	1.0690298968912066	1.0184739222129187
27.0	1.783644509662811	1.4033091947436334	127.0	1.1315182854694532	0.982837833960851
28.0	1.7183011536647503	1.4078638469179472	128.0	1.1145693346633947	1.0127228702108064
29.0	1.634750422006263	1.4355837563673655	129.0	1.144241827827199	0.9347373654445013
30.0	1.5949709119653344	1.4190023650725683	130.0	1.0344378391378803	0.9605306446552276
31.0	1.6549863064647616	1.153109993537267	131.0	1.1564309374618351	1.0179176092147828
32.0	1.575007920594592	1.1420845131079356	132.0	1.157338532402103	0.9790784870584806
33.0	1.4898351471787108	1.2615263238549232	133.0	1.1525076524095428	1.0052352527777353
34.0	1.4833270025936733	1.1709417685866357	134.0	1.0490032024961664	1.0093258847792943

Epoch	Training Loss	Validation Loss	Epoch	Training Loss	Validation Loss
35.0	1.4690176647408564	1.2161225572228431	135.0	1.130682774280247	1.0413126409053803
36.0	1.5054811225237703	1.1194933931032816	136.0	1.063908307287926	0.9462366382280986
37.0	1.4192121877033907	1.1047057057420413	137.0	1.0696223303675652	0.949197473625342
38.0	1.4142381995356172	1.0941726808746657	138.0	1.0922436875508243	0.9507547711332639
39.0	1.4184727688368999	1.1927365923921267	139.0	1.07473145302077	1.0181342482566833
40.0	1.4038214623591954	1.222982015212377	140.0	1.1385756800497384	1.0070033838351569
41.0	1.4182013561178868	1.1777732595801353	141.0	1.1898403866846758	0.9166142128407955
42.0	1.3129697075687854	1.0931795686483383	142.0	1.1718055526901008	1.0364722932378452
43.0	1.2659398409769052	1.0333578204115232	143.0	1.1308625858081014	1.0044063736995061
44.0	1.4051175933135183	1.0319288343191146	144.0	1.1199510355192916	0.9859703640143077
45.0	1.371531755012229	1.1883211170633634	145.0	1.113998552014057	0.9706973979870478
46.0	1.322748781595015	1.0030382573604584	146.0	1.1166541705790318	0.9752278099457423
47.0	1.2671883555507302	0.9623715390761693	147.0	1.1128815343897593	1.0294470166166623
48.0	1.3519595903449488	0.9837518533070883	148.0	1.0704479630392296	0.9914908106128375
49.0	1.1982765628543115	0.9929193168878555	149.0	1.162127931483258	1.005963142712911
50.0	1.281483094495042	1.1303844427069027	150.0	1.0589305576693295	0.9620907204846542
51.0	1.2783564773381204	1.0284187093377113	151.0	1.0840604585924543	1.0153198078274728
52.0	1.2837077011179208	1.0007979705929757	152.0	1.0716100862823932	0.9647033393383027
53.0	1.2833344835535925	1.0641931851704916	153.0	1.1204917351330133	0.9991291175285976
54.0	1.2358618808308042	1.0599931359291077	154.0	1.1277294489450025	1.0383158758282662
55.0	1.2283681917347407	1.1389718363682428	155.0	1.1076263808050102	0.9347692330678304
56.0	1.2353975577069853	0.9040293723344803	156.0	1.112794552558571	0.91938752780358
57.0	1.2110232035804511	0.9768775646885236	157.0	1.1339772686754401	1.0284940702219805
58.0	1.234162884939434	0.9830715343356132	158.0	1.2012257219145173	1.036897948384285
59.0	1.1884082429727216	0.9710658028721809	159.0	1.0853229789693553	0.966237464050452
60.0	1.1471294102476055	0.987448351085186	160.0	1.0474908934499985	0.9589359949032465
61.0	1.2902651030318182	1.1529646029074987	161.0	1.1012136049960788	1.0371571833888689
62.0	1.174349441620192	0.9509037946661313	162.0	1.1155880357425911	0.9955452819665273
63.0	1.1877762783402788	0.9468066131075223	163.0	1.1439217293733044	1.0057175728181997
64.0	1.2094062925281381	1.034584251542886	164.0	1.2083566517085957	0.9859377930561701
65.0	1.1898896671215395	0.9409961057205994	165.0	1.071490551520111	0.9341409911712011
66.0	1.2133106294655263	1.0627921397487323	166.0	1.0811186316878276	0.9154498333732287
67.0	1.1004479795246196	1.023066404958566	167.0	1.1872756363529908	1.0198487679163615
68.0	1.1381652501964927	1.0370575547218324	168.0	1.0293301063820832	0.9801695004105568
69.0	1.2246871320040602	1.0467761601010959	169.0	1.1057066981281554	1.0305903072158495
70.0	1.1005257754061455	1.0496965430676937	170.0	1.0270022649626087	1.0100108499328295
71.0	1.1494841737518633	0.9711947863300642	171.0	1.1057522220158935	0.9522513056794802
72.0	1.0849918922255362	0.9640756552418073	172.0	1.204476467284717	0.9816828151543935
73.0	1.084764474428686	0.9423715447386106	173.0	1.0048619138455033	0.8832322721680005
74.0	1.1948671361669563	0.9576249907414118	174.0	1.067522095726397	1.019233128676812
75.0	1.1335010130826693	1.004837842285633	175.0	1.1419250515842796	0.9776568993926048
76.0	1.1605091823783136	0.990518223742644	176.0	1.1133214789337682	1.0657138188680013
77.0	1.1193065578328039	1.0365310182174048	177.0	1.0703417679533027	1.0328313554326694
78.0	1.2189852850777763	1.0200254837671916	178.0	1.0479338415023078	1.0464870455364386
79.0	1.093446507937926	1.0197311113278071	179.0	1.218840002621475	0.9719729895393053
80.0	1.1362014433607124	1.0653843070069948	180.0	1.0458327985011546	1.0090717824796835
81.0	1.045027780437604	1.0074787711103756	181.0	1.1344553203062904	1.009123415251573
82.0	1.063627795729422	1.0402147799730301	182.0	1.1052604620169877	0.9634052241841952
83.0	1.0640427961497378	1.03180540005366	183.0	1.0507868481869984	1.0640385786692301
84.0	1.0810033561367738	1.133007578055064	184.0	1.0834719288281929	1.020915431777636

Epoch	Training Loss	Validation Loss	Epoch	Training Loss	Validation Loss
85.0	1.086573899044354	1.0266032636165618	185.0	1.1012012098068582	0.9622926702102025
86.0	1.1723823583775894	0.9424459745486577	186.0	1.10456926423244	0.989635277291139
87.0	1.0745011464433563	1.000571608543396	187.0	1.070947201628434	0.9581940993666649
88.0	1.096564537992603	0.9879064194858074	188.0	1.132206035175718	0.985691636800766
89.0	1.0930450585551728	0.9936313487589359	189.0	1.083843308227851	1.0183654750386875
90.0	1.004884030081724	1.0467615306377411	190.0	1.0739907432151468	0.928938630471627
91.0	1.0853058698033928	1.0151511569817862	191.0	1.0425955171423746	0.9160811106363932
92.0	1.1254934803547716	0.9427420968810717	192.0	1.0614205498556446	1.030308373272419
93.0	1.080010641226195	1.028528847793738	193.0	1.1129780257666917	0.9739615331093471
94.0	1.1483174742276507	1.0372228503227234	194.0	1.0653845659996335	1.0245226537187895
95.0	1.1949182692775153	1.0017878939708074	195.0	0.9902633663294906	0.8995685413479805
96.0	1.0916090208784979	1.005202337106069	196.0	1.1601626789827544	0.97757588326931
97.0	1.0636011295412715	1.0226115522285302	197.0	1.0690677904320838	0.9432818551858266
98.0	1.142116751045661	1.0320026591420173	198.0	1.0674410797935678	1.006391279399395
99.0	1.0495993151588547	1.0299364253878593	199.0	1.0990666023322515	0.9638863523801168
100.0	1.1223161732847977	1.0005273277560869	200.0	1.1375758991317642	1.0068519438306491

### Source: author statistics

#### Author contributions

Yanyu Li and Mingyi Zhao wrote the first draft, drew preliminary illustrations, and designed the research questionnaire. Mingyi Zhao conducted literature survey and analysis in the study. Jingyi Mao was deeply involved in the early stages of the survey and the digital processing of samples. Yile Chen revised the first draft, constructed the entire research idea, drew some illustrations, and translated and revised the English. Liang Zheng conducted the training of the machine learning model. Yile Chen and Lina Yan produced label samples for machine learning training during the research process. Lina Yan completed the first round of training label production. Unfortunately, the experiment did not meet expectations. Therefore, Yile Chen re-evaluated the label, produced it, and invested in machine learning training. Jingyi Mao, Yile Chen, Liang Zheng, and Mingyi Zhao participated in the revision of the manuscript.

#### Funding

This research received funding from the Guangdong Provincial Department of Education's key scientific research platforms and projects for general universities in 2023: Guangdong, Hong Kong, and Macao cultural heritage protection and innovation design Team (Funding Project Number: 2023WCXTD042) and the Fujian social science foundation project: arrangement and research on Historical Materials of A-Mazu Architectural Images in Ming and Qing Dynasties (Funding Project Number: FJ2023C053). The corresponding authors, Yile Chen and Liang Zheng, are both participating researchers in these two funded projects.

#### Data availability

The original code of the program cannot be released yet because our program is being used in other research. The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

#### Declarations

#### Competing interests

The authors declare no competing interests.

#### Institutional review board statement

Not applicable for studies not involving humans or animals.

#### Consent for publication

Not applicable for studies not involving humans.

Received: 2 January 2024 Accepted: 30 March 2024

Published online: 01 May 2024

#### References

1. The state council. circular of the state council on publishing the list of the second batch of national intangible cultural heritage and the extended list of the first batch of national intangible cultural heritage. 2011. [https://www.gov.cn/zwqk/2008-06/14/content\\_1016331.htm](https://www.gov.cn/zwqk/2008-06/14/content_1016331.htm).
2. The state council. circular of the state council on publishing the list of the third batch of national intangible cultural heritage. 2011. [https://www.gov.cn/zwqk/2011-06/09/content\\_1880635.htm](https://www.gov.cn/zwqk/2011-06/09/content_1880635.htm).
3. Wen P. (2008) investigation report of porcelain inlay handicraft workshop in chaoshan. *Zhuangshi*. 2008;02:94–6. <https://doi.org/10.16272/j.cnki.cn11-1392/j.2008.02.028>.
4. Li Y, Yin J, Wang X. (2022) a study on the technological and artistic characteristics of inlaid porcelain in traditional architectural decoration in Lingnan. *Ind Des*. 2022;05:134–6.
5. Wang Y. A study on the culture of temple roofs decoration in southern Fujian, eastern Guangdong and Taiwan. Doctoral thesis. S China Univ Technol. 2014. 1–2
6. Song Y, Liao C. Structural materials, ventilation design and architectural art of traditional buildings in Guangdong. *China Build*. 2022;12(7):900. <https://doi.org/10.3390/buildings12070900>.
7. Yuan Y, Yang Y. Analysis on the artistic expression of “flying beauty” of chaozhou inlaid porcelain. In: Yuan Y, Yang Y, editors. 7th international conference on arts design and contemporary education (ICADCE 2021). Atlantis Press; 2021. p. 121–8.
8. Cao Y, Lu Y. Analysis on porcelain inlay decoration in traditional buildings in chaozhou. In: Cao Y, Lu Y, editors. 7th international conference on arts, design and contemporary Education (ICADCE 2021). Atlantis Press; 2021. p. 111–6.
9. Zhong F. Creative transformation and innovative development of Lingnan traditional architectural culture-taking the architecture reconstruction design of liwan district in Guangzhou as an example. *J Phys Conf Ser*. 2020;1649(1):012014. <https://doi.org/10.1088/1742-6596/1649/1/012014>.



10. Jiang B, Chen S, Wang B, Luo B. MGLNN: semi-supervised learning via multiple graph cooperative learning neural networks. *Neur Netw.* 2022;153:204–14. <https://doi.org/10.1016/j.neunet.2022.05.024>.
11. 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. <https://doi.org/10.1016/j.ecoinf.2022.101919>.
12. Roy AM, Bose R, Bhaduri J. A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. *Neur Comput Appl.* 2022;34(5):3895–921. <https://doi.org/10.1007/s00521-021-06651-x>.
13. 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 Infor.* 2023;56: 102007. <https://doi.org/10.1016/j.aei.2023.102007>.
14. Wu T, Guo Y. Analysis on architectural aesthetic dimensions of the temple of the dawn complex. *IOP Conf Ser Earth Environ Sci.* 2020;567(1): 012015. <https://doi.org/10.1088/1755-1315/567/1/012015>.
15. Liu Y, Hou M, Li A, Dong Y, Xie L, Ji Y. Automatic detection of timber-cracks in wooden architectural heritage using YOLOv3 algorithm. *Int Arch Photogramm Remote Sens Spat Inf Sci.* 2020;43:1471–6. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-1471-2020>.
16. Zheng L, Chen Y, Yan L, Zhang Y. Automatic detection and recognition method of Chinese clay tiles based on YOLOv4: a case study in Macau. *Int J Archit Herit.* 2023. <https://doi.org/10.1080/15583058.2023.2246029>.
17. Li Q, Zheng L, Chen Y, Yan L, Li Y, Zhao J. Non-destructive testing research on the surface damage faced by the Shanhaiguan great wall based on machine learning. *Front Earth Sci.* 2023;11:1225585. <https://doi.org/10.3389/feart.2023.1225585>.
18. Idjaton K, Desquesnes X, Treuillet S, Brunetaud X. Transformers with YOLO network for damage detection in limestone wall images. In: Idjaton K, Desquesnes X, Treuillet S, Brunetaud X, editors. *International conference on image analysis and processing*. Cham: Springer International Publishing; 2022. p. 302–13.
19. Hu C, Dong Y, Xia G, Liu X. An automatic detection method of the mural shedding disease using YOLOv4. *Int Conf Environ Remote Sens Big Data.* 2021;12129:183–92.
20. Hou M, Hao W, Dong Y, Ji Y. A detection method for the ridge beast based on improved YOLOv3 algorithm. *Herit Sci.* 2023;11(1):167. <https://doi.org/10.1186/s40494-023-00995-4>.
21. Siountri K, Anagnostopoulos CN. The classification of cultural heritage buildings in Athens using deep learning techniques. *Heritage.* 2023;6(4):3673–705. <https://doi.org/10.3390/heritage6040195>.
22. Janković R. Machine learning models for cultural heritage image classification: comparison based on attribute selection. *Information.* 2019;11(1):12. <https://doi.org/10.3390/info11010012>.
23. Fesl J, Jelínek J, Horníčková K, Nevařilová Z, Konopa M, Feslová M. AI-based system for cultural heritage objects identification from real photos. *12th Int Conf Adv Comp Inf Technol (ACIT).* 2022. <https://doi.org/10.1109/ACIT54803.2022.9912752>.
24. Saadat, M. A., Hossain, M. S., Karim, R., & Mustafa, R. Classification of cultural heritage mosque of Bangladesh using CNN and Keras model. In *Intelligent computing and optimization: proceedings of the 3rd International Conference On Intelligent Computing And Optimization 2020*. 2021. 647–658.
25. Xiong Y, Chen Q, Zhu M, Zhang Y, Huang K. Accurate detection of historical buildings using aerial photographs and deep transfer learning. *Int Geosci Remote Sens Symp.* 2020. <https://doi.org/10.1109/IGARSS39084.2020.9323541>.
26. Zou H, Ge J, Liu R, He L. Feature recognition of regional architecture forms based on machine learning: a case study of architecture heritage in Hubei province. *China Sustainability.* 2023;15(4):3504. <https://doi.org/10.3390/su15043504>.
27. Girsang ND. Literature study of convolutional neural network algorithm for batik classification. *Brill Res Artif Intell.* 2021;1(1):1–7. <https://doi.org/10.47709/brilliance.v1i1.1069>.
28. Horn C, Ivarsson O, Lindhe C, Potter R, Green A, Ling J. Artificial intelligence, 3D documentation, and rock art—approaching and reflecting on the automation of identification and classification of rock art images. *J Archaeol Method Theor.* 2022;29(1):188–213. <https://doi.org/10.1007/s10816-021-09518-6>.
29. Liu E. Research on image recognition of intangible cultural heritage based on CNN and wireless network. *EURASIP J Wirel Commun Netw.* 2020;2020:1–12. <https://doi.org/10.1186/s13638-020-01859-2>.
30. Guo X. Rooftop with XICHU—a study of porcelain carving in Fujian, Guangdong and Taiwan since the Qing dynasty. *Fujian Norm Univ.* 2021. <https://doi.org/10.27019/d.cnki.gfjsu.2021.000006>.
31. Xue Y. Modern Lingnan architectural decoration research. Doctoral thesis. *S China Univ Technol.* 2012. 161–166.
32. Yakovlev A, Lisovychenko O. An approach for image annotation automatization for artificial intelligence models learning (Підхід до автоматизації анотування зображень для навчання моделей штучного інтелекту). *Адаптивні системи автоматичного управління.* 2020;1(36):32–40. <https://doi.org/10.20535/1560-8956.36.2020.209755>.
33. Redmon J, Farhadi A. Yolov3: an incremental improvement. *arXiv preprint arXiv 1804.2018;02767:1–6.* <https://doi.org/10.48550/arXiv.1804.02767>.
34. Wang ZZ, Xie K, Zhang XY, Chen HQ, Wen C, He JB. Small-object detection based on yolo and dense block via image super-resolution. *IEEE Access.* 2021;9:56416–29. <https://doi.org/10.1109/ACCESS.2021.3072211>.
35. Yan L, Chen Y, Zheng L, Zhang Y. Application of computer vision technology in surface damage detection and analysis of shedthin tiles in China: a case study of the classical gardens of Suzhou. *Herit Sci.* 2024;12(1):72. <https://doi.org/10.1186/s40494-024-01185-6>.
36. Terven J, Cordova-Esparza D. A comprehensive review of YOLO: from YOLOv1 to YOLOv8 and beyond. *arXiv Preprint.* 2023. <https://doi.org/10.48550/arXiv.2304.00501>.
37. Terven J, Córdoba-Esparza DM, Romero-González JA. A comprehensive review of yolo architectures in computer vision: from yolov1 to yolov8 and yolo-nas. *Mach Learn Knowl Ext.* 2023;5(4):1680–716. <https://doi.org/10.3390/make5040083>.
38. Xu N. Inheritance vein and cultural implication of the intangible cultural heritages porcelain inlay art in chaoshan. Master's thesis. *Guangdong Univ Technol.* 2014. 2.

## Publisher's Note

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