# Construction and Retrieval of Shu Brocade Pattern Image Resources for Digital Preservation

**Xuan Lu[1], Zhicong Rui[2], Xingren Wang[1]**
*[1]Key Laboratory of Digital Preservation and Transmission of Weaving and Embroidery Heritage, Sichuan Provincial Department of Culture and Tourism (Chengdu Qinggong Polytechnic University), Sichuan Provincial Department of Culture and Tourism, Chengdu, Sichuan, China*
*[2]Sichuan Engineering Research Center for Functional Development and Application of High-Performance Specialty Textile Materials, Chengdu, Sichuan, China*

**Abstract: The government currently places great emphasis on the digital preservation of intangible cultural heritage, and given that Shu brocade is renowned throughout the country, the digital archiving of its patterns is of utmost importance. Addressing issues such as the lack of domain-specific image classification models for Shu brocade patterns, scarcity of data, and insufficient digital preservation of related cultural heritage, this study established a software-based digital platform using over 230 images of Shu brocade patterns as a foundation. At the same time, it carried out research on resource indexing and pattern classification uses for the Shu brocade pattern database. It built a searchable pattern database. This database arranges images by names, descriptions, themes, and other metadata. It also allows keyword searches and category-based filtering. To develop a lightweight Shu Brocade model, it used the ImageNet-pretrained Swin Transformer Tiny model. This model carried out classification analysis on a small dataset. The dataset had 113 images. The chosen model reached a validation accuracy of 85.71%. Its test accuracy was 53.33%. These results gave preliminary findings for the classification of Shu brocade patterns. The results indicate that the lack of category samples and distribution imbalances limited the accuracy of classification predictions. Further training requires more images and better coverage of all classes.**

**Keywords: Shu Brocade Patterns; Digital Preservation; Resource Organization; Image Classification and Retrieval**

## 1. Introduction

Shu brocade patterns convey information about subject matter, composition, color, and weaving culture. Their digital preservation concerns both the retention of visual materials and their subsequent use in research and design. As image collections grow, folder-based storage and manual browsing do not readily support searches by name, subject, or visual content. When image acquisition precedes cataloging, the quantity of images and the completeness of their descriptions may develop at different rates. Images can therefore be preserved without being discoverable through the terms familiar to their users.

## 2. Related Work

Research on cultural heritage digitization has expanded from acquisition and archiving to resource management and knowledge use. In a bibliometric analysis of studies from 2019 to 2024, Li et al. identified continuing applications of digital technologies in heritage management, presentation, and interdisciplinary research, alongside constraints involving technological suitability and data protection [1]. In terms of relevant research on traditional pattern processing, Quan et al. create a combination between the Miao batik knowledge atlas and an improved version of ResNet34 classifier, which aims to associate patterns with cultural information through text queries and image classification [2]. With regard to correlative research of the Shu brocade, Wang's team adopted the super-resolution reconstruction and vector drawing technology in order to optimize the digital representation of historical patterns [3]. These aforementioned studies belongs to the efforts conducted by many researchers apply deep learning and digital technology to empower the preservation of traditional Shu Brocade patterns. Although, this work also involves

 Copyright @ STEMM Institute Press

multiple other fields, such as digitalization of intangible cultural heritage, machine vision, and image processing technology.

Take image retrieval technology as an example, it usually relies on textual descriptions or visual features. The gap between visual features and user semantic interpretation remains a fundamental issue in content-based image retrieval today [4]. To tackle this issue, Sun et al. attempted to address the recognize patterns on Nantong blue printed fabric by establishing relevant databases and assigning corresponding category labels to achieve better results. The tests operated on serveral models such as SSD, Faster R-CNN, and YOLOv8. The experimental data acquired from those tests demonstrates the significance of database organization in establishing connections with visual recognition [5]. In addition, Varshney et al. [6] proposed a joint method of InceptionV3 and InceptionResNetV2 for the retrieval task of traditional Indian textile patterns. This method integrates the deep features of these two networks and combines them with principal component analysis, clustering, and correlation feedback information to achieve retrieval optimization. These methods focus respectively on pattern classification and visual similarity retrieval based on image content, and they rely heavily on the accuracy of data descriptions and manually assigned labels. Swin Transformer offers an alternative approach to image representation through hierarchical features and shifted-window attention [7]. For a small collection of Shu brocade images, classification outputs need to be verified before they can be used as category metadata for retrieval.

This paper examines the relationship among the development of a Shu brocade pattern image database, image retrieval, and domain-specific image classification. It examines the resource development and access for the Shu brocade pattern database, as well as the predictive capabilities of the model under conditions of limited sample categories. Building on existing database records and retrieval functions, the paper analyzes the development of a domain-specific classifier for Shu brocade patterns by incorporating training logs and test results from the Swin-T network.

## 3. Resource Organization and Pattern Classification

As shown in Figure 1, the resource database platform we developed contains more than 230 images of Shu brocade patterns, featuring designs with various symbolic meanings, from different dynasties, and in various styles, all of which have been uniformly categorized and labeled using simple tags.

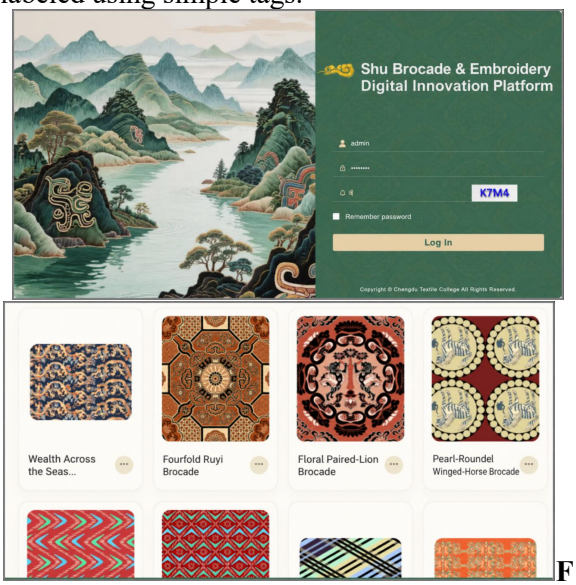



**Figure 1. Platform Interface and Digital Pattern Display**

### 3.1 Data and Analytical Framework

The image classification dataset used in this experiment consists of 113 images, organized by category into training, validation, and test sets. The database contains more than 230 images. The classification experiment uses 113 images, divided into 77 training images, 21 validation images, and 15 test images. The name records in the image database serve a different purpose than the dataset labels used in the classification experiments: the former facilitates the identification and indexing of image resources, while the latter provides the groundwork for model training and result evaluation.

The process structure for building an image resource database and using it to assist large-scale model classification is shown as Figure 2. The structure shown on the left side of the figure represents the image and its related label information. Users can retrieve these images by pattern name or by label category. On the right, the classifier has four output classes: Animal, Flower, Human, and Other. However, the training set contains examples from only Animal, Flower, and Other. The model then outputs a predicted category. The original labels for the images on both sides were created manually. This approach separates the pattern database

resources from model inference, allowing the final classification results output by the large model to be verified before the search tags are generated. The dashed line indicates the manually organized classification results, which are ultimately imported into the database.

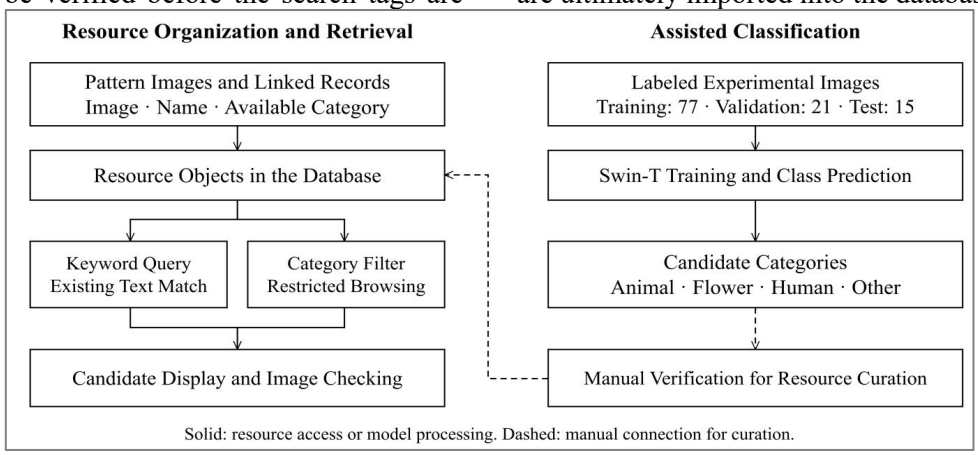



**Figure 2. Resource Organization and Assisted Classification**

### 3.2 Image Records and Resource Indexing

Database organization centers on each pattern image and its associated access descriptions. The relational model supports information organization and use through explicit data relationships [8]. The emphasis on findability and reusability in the FAIR principles also suggests considering descriptions together with access requirements [9]. Existing resource records are therefore described in terms of three elements: the image, its name or textual description, and its available category. The image provides visual evidence, the name provides a textual search cue, and the category defines a browsing group. This is a logical description of resource relationships and does not assume a particular database product or physical index structure.

Resource records must accommodate missing information. An image without a category remains available for browsing and should not automatically be assigned to the other class. Other is an experimental label represented by directory samples; it differs from an unannotated status. Similarly, a name can identify an image but cannot replace historical verification of its date, technique, or cultural meaning. Retaining unknown information as pending curation helps prevent missing values from being interpreted as confirmed semantics.

Keyword queries and category filters impose different access conditions. Keyword queries match user input against existing searchable text and suit users who know a name or a subject-related term. Category filters return records associated with an available category and support exploratory browsing when the exact name is unknown. The contribution of category metadata to image browsing has been discussed in previous research [10]. In this application, both routes lead to resource records, while image displays allow users to verify the results. A textual match or category agreement therefore identifies a candidate resource rather than guaranteeing its relevance.

Retrieval results depend on descriptive coverage. A composite pattern may contain an animal motif surrounded by floral borders. If its name describes only the central animal, a flower-related query may not return the record. A missing result can thus reflect either an absent resource or an incomplete description. Category filters likewise cover only records with established category associations. Direct image browsing should consequently coexist with textual access, and an analysis of retrieval must distinguish executable query functions from the content actually covered by the available data.

### 3.3 Candidate Category Prediction Using Swin-T

The classifier is implemented with PyTorch and timm using the Swin Transformer Tiny architecture. RGB images of 224 × 224 pixels are supplied as input, and locally stored ImageNet-pretrained weights are loaded before training. Swin-T divides each image into 4 × 4 patches and applies linear embedding before four feature-extraction stages. The stages contain 2, 2, 6, and 2 blocks, with 96, 192, 384, and 768 channels, respectively. Window-based and shifted-window attention alternate within the network. Patch merging reduces spatial

 Copyright @ STEMM Institute Press

resolution between stages, producing a hierarchical image representation [7].

The Swin-T classification structure proposed in this article is shown in Figure 3. Normalize and globally average pool after Stage 4, and output the generated 768 dimensional feature vectors to the linear classification head. The classification header outputs logits for four categories: animals, flowers, humans, and others. Finally, Softmax converts logits into classification prediction probabilities. Then use the class with the highest probability value as the prediction result.

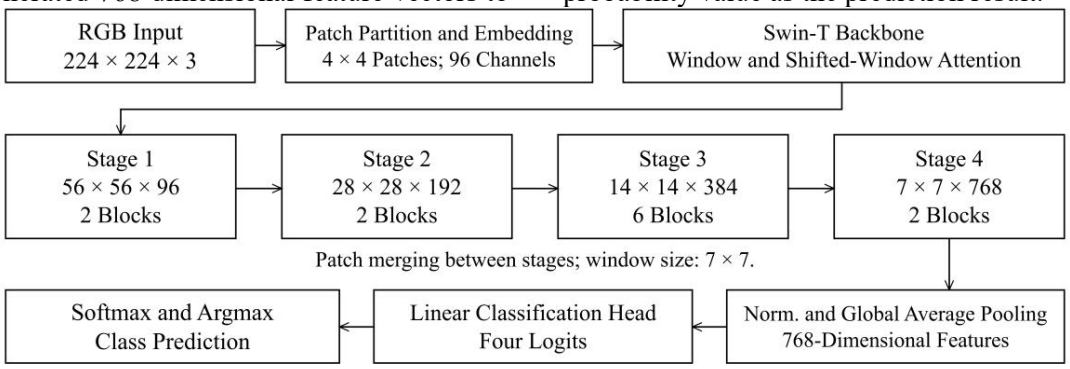



**Figure 3. Swin-T Architecture for Pattern Classification**

After the output of the classification model, resource management is also required, which includes two independent steps: "generating candidate labels" and "updating database records". But before that, manual review of the original images is required, and it must be confirmed that the final classification results match the original labels in the database before these results can be used as the correct classification labels. However, for images with unclear themes, multiple element types, and insufficient confidence scores, the model keeps them in an unassigned label state. In order to verify the accuracy of these images and reference labels through manual calculation in subsequent validation.

## 4. Experiments and Results

### 4.1 Data Splits and Evaluation

The dataset used for the experiment mainly includes four categories: animals, flowers, humans, and others. The floral category also includes some plant patterns. As shown in Figure 4, there are partial samples of four categories, where the segmented training set, validation set, and test set contain 77, 21, and 15 images, respectively. The distribution of images for each category is shown in Table 1.

During model training, four labels will be periodically output at each epoch, with Batch_2 set to 16 and training configured for 30 epochs. The AdamW optimizer was used with an initial learning rate of $1 \times 10^{-4}$, with the learning rate scheduled using Cosine Annealing LR. During the training process, the validation set is periodically called to evaluate the performance of the check_ point node model and store check_ point related data. After training, test sets are called to verify the classification effect and evaluate the accuracy of classification. The overall performance indicator is defined as the predicted confidence value of a single image.

### 4.2 Database Management Platform

As shown in Figure 5, the database management platform interface displays the Shu brocade patterns stored in the library in the form of cards, and has a search function to retrieve corresponding resources based on pattern names and other information. In addition, this platform also includes functions such as pattern modeling, repair, recognition, etc., and comes with a management subsystem.



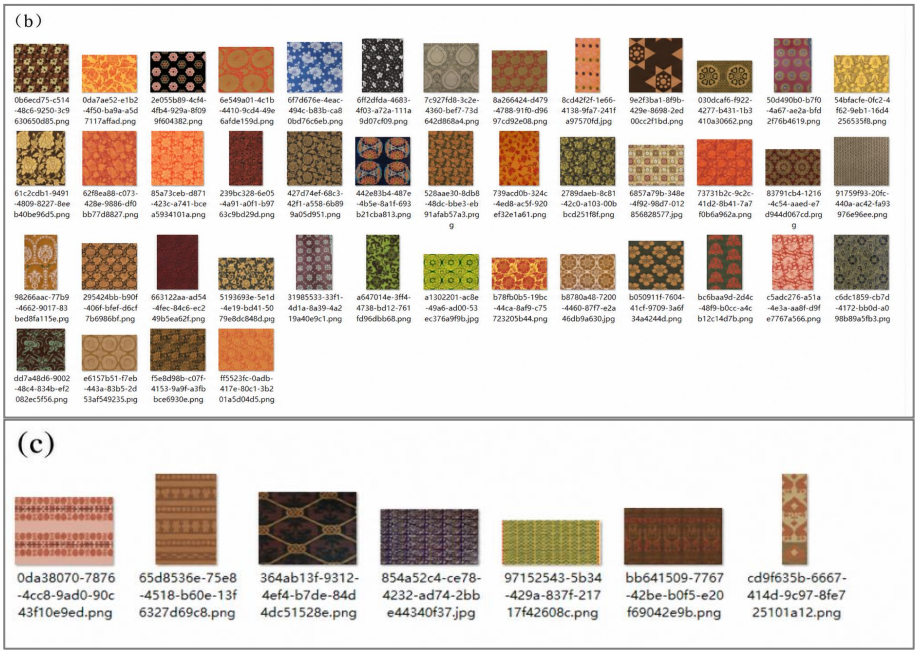



**Figure 4. Partial Datasets Samples: (a) Animal; (b) Flower; (c) Other**

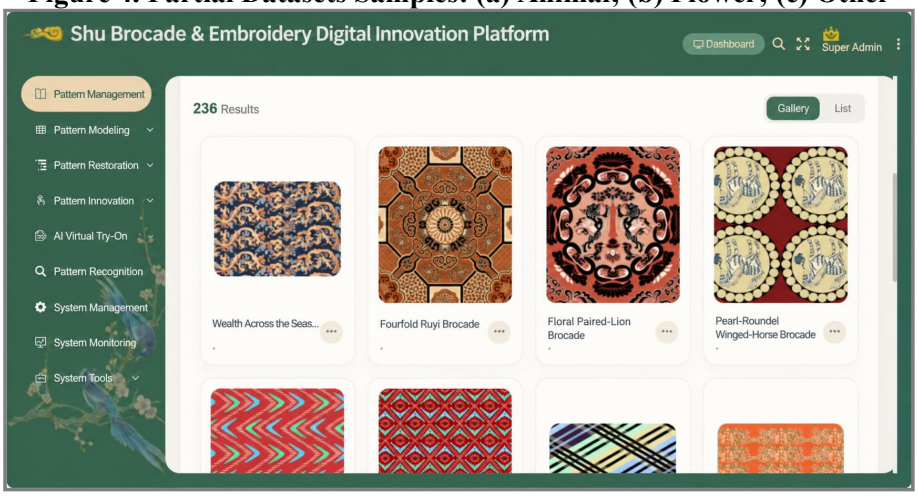



**Figure 5. Linked Display of Pattern Images and Names**

**Table 1. Class Distribution of the Pattern Dataset**

| Class | Training | Validation | Test | Total |
|---|---|---|---|---|
| Animal | 27 | 7 | 5 | 39 |
| Flower | 43 | 12 | 7 | 62 |
| Human | 0 | 0 | 1 | 1 |
| Other | 7 | 2 | 2 | 11 |
| Total | 77 | 21 | 15 | 113 |

Records with informative names can be accessed through textual cues, while records with categories can be included in corresponding browsing groups. These routes support targeted searches and exploration within a restricted set. Direct browsing keeps incompletely described images visible and available for later cataloging. The platform thus connects images with access cues. Whether this reduces search time remains to be evaluated through controlled user studies using the same retrieval tasks.

### 4.3 Training Observations

Figure 6 presents the logs for the first five training epochs. Validation accuracies are 61.90%, 85.71%, 71.43%, 71.43%, and 71.43%, with the highest displayed value occurring at epoch 2. The validation set contains only 21 images, so each additional correct prediction changes accuracy by approximately 4.76 percentage points. Accuracies of 85.71% and 71.43% correspond to 18/21 and 15/21 correct predictions, respectively, a difference of three images. These fluctuations must be interpreted in light of the small sample size; a single high value does not establish stable convergence.

Training accuracy increases from 51.95% at epoch 1 to 98.70% at epoch 4 and is 96.10% at epoch 5, without a corresponding increase in

 Copyright @ STEMM Institute Press

validation accuracy. This gap shows that we need to consider training fit and validation variability when data is limited. The screenshot only covers early training. Therefore, it cannot determine the best checkpoint among all completed epochs. It also cannot show later performance trends. We therefore report 85.71% as the highest validation accuracy that can be seen in the displayed interval.

### 4.4 Test Results and Class Differences

The individual prediction in Figure 7 has a confidence score of 0.9677, or approximately 96.77%; this is not validation accuracy. The selected model correctly classifies 8 of the 15 test images, yielding an overall test accuracy of 53.33%. Table 2 summarizes the results by class. Four Animal images and four Flower images are classified correctly, giving recalls of 80.00% and 57.14%, respectively. Recall is zero for Human and Other. The mean of the four class recalls is 34.29%, below overall accuracy, indicating that correct predictions are concentrated in Animal and Flower and that performance is uneven across classes.

The weights used for testing correspond to the checkpoint saved with a validation accuracy of 85.71% in Figure 6. Validation data inform checkpoint selection during training, whereas the test set evaluates the selected weights. Test performance is lower than validation performance. The sets differ in size and class distribution, and Human appears only in the test set. This difference should therefore be interpreted in relation to sample composition and model selection, rather than attributed solely to one architectural factor.

```
(.venv) PS D:\01_work\01_develop\01_code\19_cursor\ai\sjsx> python train.py
Using device: cuda
Detected 4 classes: ['animal', 'flower', 'human', 'other']
Train samples: 77
Val samples: 21
Loaded local pretrained weights: weights/model.safetensors
Epoch [1/30] train_loss=1.0347 train_acc=0.5195 val_loss=0.7566 val_acc=0.6190
Saved best model to weights\best.pth (val_acc=0.6190)
Epoch [2/30] train_loss=0.6041 train_acc=0.7403 val_loss=0.6346 val_acc=0.8571
Saved best model to weights\best.pth (val_acc=0.8571)
Epoch [3/30] train_loss=0.4685 train_acc=0.8052 val_loss=0.7459 val_acc=0.7143
Epoch [4/30] train_loss=0.2135 train_acc=0.9870 val_loss=0.8014 val_acc=0.7143
Epoch [5/30] train_loss=0.1916 train_acc=0.9610 val_loss=0.9018 val_acc=0.7143
```

**Figure 6. Training Logs for the First Five Swin-T Epochs**

```
(.venv) PS D:\01_work\01_develop\01_code\19_cursor\ai\sjsx> python predict.py "dataset/test/flower/3773ea52-e18b-4d30-9fad-ebc01709
image_path: D:\01_work\01_develop\01_code\19_cursor\ai\sjsx\dataset\test\flower\3773ea52-e18b-4d30-9fad-ebc017098c98.png
predicted_class: flower
chinese_label: flower
confidence: 0.9677
```

**Figure 7. Classification Prediction for an Individual Pattern Image**

**Table 2. Classification Results on the Test Set**

| Class | Test Images | Correct Predictions | Recall (%) |
|---|---|---|---|
| Animal | 5 | 4 | 80.00 |
| Flower | 7 | 4 | 57.14 |
| Human | 1 | 0 | 0.00 |
| Other | 2 | 0 | 0.00 |

There is an efficient approach in evaluation that only calculates baseline accuracy based on category distribution, which can be used as a reference. This method does not require the use of image visual features or additional model experiments, thus saving computational costs. The specific approach of this method is to predict each test image as the category with the most samples in the training set, namely Flower. According to the original data, this will correctly classify 7 Flower images; If the test set consists of 15 images, the reference accuracy is 7/15=46.67%. However, in reality, the evaluated model only correctly judged one image more than this reference baseline. This indicates that looking solely at the overall accuracy does not necessarily mean that the model has achieved stable and actual benefits from visual feature learning. If the results of the four categories are viewed in order by category, more information can be observed, and it can be seen which categories' performance supports the overall

Copyright @ STEMM Institute Press 

performance.

But there is still a problem, that is, the training set does not contain human images; As a result, the model also misclassifies the only human image in the test set. In addition, the other class has only seven training images. It also has only two test images. This small number of training samples may affect the classifier's performance. Furthermore, we cannot tell exactly which categories received these two types of errors. The test set is small overall. In this test set, one image is about 6.67 percentage points. Because of this, our analysis focuses on explaining the classification patterns we observed in the existing samples. We do not extend the results to represent the classification accuracy of the entire Shu brocade resource repository. So we do not treat the classification results for the "Human" category as representative.

## 5. Discussion

The descriptive information associated with the images determines the search methods on the Shu Brocade Pattern Platform. For small-scale pattern databases, we should first refine names and categories. This work gives a reliable basis for keyword searches and categorical filtering. In addition, people should manually verify predicted labels before they add them to database records. In this way, classification errors do not affect search results. We should also distinguish “Other” from “Unlabeled.” The former is a specific classification label. The latter means that a determination has not yet been made. If people use them interchangeably, the consistency of the database resources would be weakened.

In the testing experiments, the model lacked training samples for the “ Human ” category. There are also few samples from other categories. These issues limit the classification ability of the model.

In future work, we will focus on collecting more samples and building validation labels. Before segmenting data, duplicate and nearly duplicate images should be manually identified to avoid overly similar image frames, thereby preventing related images from appearing in different subsets. In addition, multiple repeated experiments are required to evaluate the stability of the results.

## 6. Conclusion

The aim of this study is to establish a mature database of Sichuan brocade pattern images, and to investigate data retrieval and classification methods for lightweight models of Sichuan brocade patterns based on the database. This study is based on Swin-T to construct a classifier structure model for Sichuan brocade. In the first five recorded training cycles, the highest validation accuracy reached 85.71%. The testing accuracy of the selected model is 53.33%. Its macro average recall rate is 34.29%.

The problem is that this model performs poorly in the categories of "human" and "other". This is mainly due to the small sample size and incomplete class coverage of this category. Moreover, the importance of manually verifying the accuracy of classification predictions before label assignment operations was also verified during training. The results indicate that the digital preservation of Shu brocade patterns requires support in terms of data quality and quantity. Classification models play a crucial role in the recognition and retrieval of Shu brocade images, while also considering the model's ability to locate resources and the reliability of its annotations.

## Acknowledgments

This work was supported by the Key Laboratory of Digital Preservation and Transmission of Weaving and Embroidery Heritage, (Chengdu Qinggong Polytechnic University), Sichuan Provincial Department of Culture and Tourism.) (2025SYSZC03); the Key Laboratory of Digital Preservation and Transmission of Weaving and Embroidery Heritage, (Chengdu Qinggong Polytechnic University), Sichuan Provincial Department of Culture and Tourism.) (2024SYSZD04)

## References

[1] LI W, XIE Q, AO J, et al. Systematic review: a scientometric analysis of the status, trends and challenges in the application of digital technology to cultural heritage conservation (2019–2024). npj Heritage Science, 2025, 13: 90. DOI: 10.1038/s40494-025-01636-8.

[2] QUAN H, LI Y, LIU D, et al. Protection of Guizhou Miao batik culture based on knowledge graph and deep learning. Heritage Science, 2024, 12: 202. DOI: 10.1186/s40494-024-01317-y.

[3] WANG W J, LIU Y, XIAO L, et al. Digitalization of ancient Shu brocade patterns based on an improved generative

 Copyright @ STEMM Institute Press

adversarial network and vector drawing technology. Journal of Silk, 2023, 60(11): 18-27.

[4] SMEULDERS A W M, WORRING M, SANTINI S, et al. Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(12): 1349-1380. DOI: 10.1109/34.895972.

[5] SUN K K, HUANG J W, YUAN Y Y, et al. Classification and recognition of the Nantong blue calico pattern based on deep learning. Journal of Engineered Fibers and Fabrics, 2024, 19: 15589250241270618. DOI: 10.1177/15589250241270618.

[6] VARSHNEY S, SINGH S, LAKSHMI C V, et al. Content-based image retrieval of Indian traditional textile motifs using deep feature fusion. Scientific Reports, 2024, 14: 10035. DOI: 10.1038/s41598-024-56465-9.

[7] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical vision transformer using shifted windows//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 10012-10022. DOI: 10.1109/ICCV48922.2021.00986.

[8] CODD E F. A relational model of data for large shared data banks. Communications of the ACM, 1970, 13(6): 377-387. DOI: 10.1145/362384.362685.

[9] WILKINSON M D, DUMONTIER M, AALBERSBERG I J, et al. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 2016, 3: 160018. DOI: 10.1038/sdata201618.

[10] YEE K P, SWEARINGEN K, LI K, et al. Faceted metadata for image search and browsing//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM, 2003: 401-408. DOI: 10.1145/642611.642681.

Copyright @ STEMM Institute Press 