IDENTIFICATION OF MIAO EMBROIDERY IN SOUTHEAST GUIZHOU PROVINCE OF CHINA BASED ON CONVOLUTION NEURAL NETWORK

Chune ZHANG^{1,2}, Song WU³, Jianhui CHEN^{1,*}

¹ Fashion & Art Design Institute, Donghua University, Shanghai, 200051, China;

² College of Sericulture, Textile and Biomass Sciences, Southwest University, Chongqing, 400715, China;

³ College of Computer and Information Science, Southwest University, Chongqing, 400715, China

*Corresponding author e-mail: chenchen@dhu.edu.cn

Abstract:

Miao embroidery of the southeast area of Guizhou province in China is a kind of precious intangible cultural heritage, as well as national costume handicrafts and textiles, with delicate patterns that require exquisite workmanship. There are various skills to make Miao embroidery; therefore, it is difficult to distinguish the categories of Miao embroidery if there is a lack of sufficient knowledge about it. Furthermore, the identification of Miao embroidery based on existing manual methods is relatively low and inefficient. Thus, in this work, a novel method is proposed to identify different categories of Miao embroidery by using deep convolutional neural networks (CNNs). Firstly, we established a Miao embroidery image database and manually assigned an accurate category label of Miao embroidery to each image. Then, a pre-trained deep CNN model is fine-tuned based on the established database to learning a more robust deep model to identify the types of Miao embroidery. To evaluate the performance of the proposed deep model for the application of Miao embroidery categories recognition, three traditional non-deep methods, that is, bag-of-words (BoW), Fisher vector (FV), and vector of locally aggregated descriptors (VLAD) are employed and compared in the experiment. The experimental results demonstrate that the proposed deep CNN model outperforms the compared three non-deep methods and achieved a recognition accuracy of 98.88%. To our best knowledge, this is the first one to apply CNNs on the application of Miao embroidery categories recognition. Moreover, the effectiveness of our proposed method illustrates that the CNN-based approach might be a promising strategy for the discrimination and identification of different other embroidery and national costume patterns.

Keywords:

Southeast Guizhou in China; Miao embroidery identification; Deep CNNs

1. Introduction

Miao is one of the most populous minority in China, distributed over different regions of Guizhou Province. Miao people do not have their own language; their national history is partly recorded in embroidery patterns, which decorate their costumes, so the embroidery not only has a decorative function but also a recording function [1]. In the Guizhou province, the southeast area has the most Miao nationalities, and here, the most exquisite and abundant embroideries can be found [2]. Therefore, a research of Miao embroidery in the southeast Guizhou will be represented. Miao embroidery in the southeast area is a kind of precious intangible cultural heritage [3], an important part of Chinese-minority costume culture, and also a national costume handicrafts and textiles. As shown in Fig.1, the embroideries usually consist of decorations on different parts of clothing such as neckline, shoulders, and sleeves (Fig.1). Particularly, all "Miao embroidery" in this paper refer to the Miao embroidery in the southeast Guizhou of China.

Now, the current situation of keeping Miao embroidery is problematic: on the one hand, the older generation with embroidery production skills gradually passed away; on the other hand, the younger generation is not sufficiently motivated

to learn the Miao embroidery skills. These lead to a crisis for the survival of Miao embroidery techniques. Furthermore, Miao embroidery is sold as a commodity by many vendors in the market [4-6], leading to a serious decrease in the number of existing embroideries. Therefore, researches on embroidery protection is necessary. The Miao people in southeast Guizhou have different branches, where each branch uses different embroidery techniques. The embroidery techniques include 12 categories: wring stitch, barbola embroidery, wrinkling stitch, split line stitch, warp-width counting stitch, and so on [7]. It is difficult to distinguish the difference without sufficient knowledge of Miao embroidery. Yet, the identification of Miao embroidery is the premise for further research. Based on the existing manual identification methods, which need a large amount of embroidery knowledge, it cannot satisfy the requirement of efficiency and accuracy of Miao embroidery identification. Quickly and accurately identifying and classifying the categories of Miao embroidery becomes a challenging research.

Inspired by the recent computer vision technologies that have been widely used to recognize the types of clothes, we proposed to use computer vision technologies to identify the categories of Miao embroidery. A traditional image recognition system

Figure 1. Examples of Miao embroidery and clothing of Southeast Guizhou: (a) a single embroidery (b) decorated sleeve part (c) embroidered clothing

based on visual word representations mainly owes its success to locally invariant descriptors and large visual codebooks. BoW [8] encodes local descriptors into a histogram according to the occurrence frequency of each visual word. FV [9] uses a Gaussian mixture model (GMM) to cluster visual words where the gradients of visual words corresponding to particular parameters in GMM are summed and concatenated into an FV image representation. VLAD [10] accumulates the difference of each local descriptor to the visual words and concatenates these accumulated values to represent an image.

In recent years, deep learning techniques, particularly, convolutional neural networks (CNNs) have achieved great success in various computer vision applications. It has significantly powerful feature representations abstracting and learning capabilities with a high performance in large-scale image pattern recognition [11]. CNNs have attracted a lot of attention, with the aim at developing an efficient, automated, and accurate system for image identification [12]. In particular, CNNs have the powerful learning and abstracting capability of efficient feature representations [13]; the CNNs models have been widely used in agriculture, medical care, education, energy, industrial inspection, and other fields. In the field of culture objects, Jia Xiaojun et al. [14] proposed a classification method based on CNNs for the image elements constituting the blue calico vein pattern; the experimental results show that average accuracy of 99.61% and detection accuracy of 98.5%. Sheng Jiachuan et al. [15] proposed an algorithm of Chinese paintings sentiment recognition via CNN optimization with human cognition. Li Rongrui et al. [16] proposed a CNN model for the recognition of traditional headdress images and the experiment results with a recognition accuracy of 96.25%. Ai Hu et al. [17] proposed an effective dialect identification model for Guizhou dialect identification based on CNNs. In the field of textile and garments, Wu Huan et al. [18] proposed a classification approach of clothing silhouettes based on the CaffeNet model of CNNs. The experimental results show that the novel approach can classify the silhouette of women's trousers, with a classification accuracy of up to 95%. Wang Fei et al. [19] proposed a novel identification method based on CNNs to sufficiently identify the cashmere and wool fiber

images. Liu Zhengdong et al. [20] have used CNNs to recognize various shapes and sizes of suits and achieved an accuracy of over 90%. He Xiaoyun et al. [21] have used RCNN deep learning neural networks to detect foreign fibers in seed cotton, and the detecting accuracy is 90.3%. Wang Wenwen et al. [22] proposed an identification method based on CNNs to identify the position of spinning yarn breakage, with an accuracy rate of over 97%. Wang Shanna et al. [23] proposed a method for fabric image motion classification based on CNNs for evaluating necktie patterns. Corona et al. [24] used a hierarchy of three CNNs with different levels of specialization to identify the type of garment. Zhao Yudi et al. [25] and Jing Junfeng et al. [26] have used CNNs for automatic fabric defect detection.

To our knowledge, this is the first paper to apply the deep CNN models on the recognition of Miao embroidery; there have been no reports on the identification of Miao embroidery based on CNNs. Because of the visual importance of exterior for the identification of Miao embroidery, Miao embroidery images can be used for the discrimination of Miao embroidery using deep learning methods. Therefore, we propose an image classification algorithm based on the Inception-v4 [27] model. The experiments are conducted on our database of Miao embroidery images, and the results show the effectiveness of this approach for the classification of Miao embroidery. At the same time, we compare our method with traditional classification models BoW, FV and VLAD. The results indicate that the classification based on CNN (Inception-v4) achieves the best results. As an artificial intelligence technology, the application of deep learning to the classification of Miao embroidery to some extent helps with the protection of this intangible cultural heritage.

2. Materials and methods

2.1. Data acquisition

To have a deep study of Miao embroidery for further intangible cultural heritages protection, we establish a database with the collected images by field investigation. Most of the Miao

Figure 2. Different types of Miao embroidery

Table 1. Details of each category in the Miao embroidery image database

Different species of Miao embroidery	Number
Wring stitch	455
barbola embroidery	305
Wrinkling stitch	245
Split line stitch	455
warp-width counting stitch	245

people in southeast area of Guizhou Province live in cottages that are located in geographical remote villages, making it difficult to collect image samples. The embroidery images were collected by field investigation during five trips: January 2016, March 2016, March 2017, November 2017, and April 2019. During these investigations, we went to different villages and communicated with local residents to understand and collect different types of embroidery. Based on these useful data and knowledge, five types of Miao embroidery (a total number of 1,705 of Miao embroidery images) were collected for the database established, as well as different kinds of research. To our best knowledge, this is the first one to establish an accurately labeled image database of Miao embroidery. Some examples of established Miao embroidery database are shown in Fig. 2 (i.e., wring stitch (Fig. 2a), barbola embroidery (Fig. 2b), wrinkling stitch (Fig. 2c), split line stitch (Fig. 2d), and warp-width stitch (Fig. 2e)). Meanwhile, the details about the different species of Miao embroidery are presented in Table 1.

2.2. Visual words-based Miao Embroidery identification

2.2.1. Bag-of-Words (BoW)

The BoW model is inspired from simple document retrieval systems. In the BoW model, salient points are first detected from each image in the training dataset. Then, each salient point is represented as a local descriptor. These local descriptors are then clustered into a vocabulary of visual words. Therefore, an image is finally represented as a histogram over a set of learned visual words after quantizing each of the local descriptors to the nearest visual word.

2.2.2 Fisher Vector (FV)

The FV model first extracts salient local descriptors from the training dataset and employs the GMM to estimate a parametric probability distribution over the extracted local descriptor space. The extracted local descriptors are then assumed to be sampled independently from this probability distribution. Each local descriptor is represented by the gradient of the probability distribution at that feature with respect to its parameters. Gradients corresponding to all the features with respect to a particular parameter are summed and concatenated to generate the final FV representation.

2.2.3 A Vector of Locally Aggregated Descriptors (VLAD)

VLAD is similar to the BoW model. A visual words vocabulary is first learned via a k-means cluster. Each local descriptor is associated to its nearest visual word in the vocabulary. The

Figure 3. Miao embroidery classification based on the BoW model

Figure 4. Miao embroidery classification model based on FV

Figure 5. Miao embroidery classification model based on VLAD

VLAD representation is to accumulate the residuals belonging to each of the visual words.

2.3. Deep CNN-based Miao embroidery identification

Generally, the CNN is a hierarchical architecture. It consists of several stacked convolutional layers, which are optionally followed by a normalization layer and a spatial pooling layer, fully connected layers, and a loss function on top. The convolutional layers generate feature maps by linear convolutional filters followed by nonlinear activation functions. The fully connected layer has full connections to all activations in the feature maps to generate a one-dimensional vector. The one-dimensional feature vector is then fed into the loss function for optimization. A general CNN architecture is shown as Fig. 6.

There are two main stages for training the convolutional neural network: a forward stage and a backward stage. The forward stage is to represent the input image with the current parameters (weights and bias) in each layer. Then, the output from the last layer is used to compute the loss function with

the ground truth labels. Based on the loss cost, the backward stage computes the gradients of each parameter with chain rules. All the parameters are updated based on the gradients and are prepared for the next forward computation. After sufficient iterations of the forward and backward stages, the network could be optimized.

2.3.1 Model architecture

For Miao embroidery classification, the classic Inception-v4 deep CNN model is employed and fine-tuned to achieve high-performance Miao embroidery classification.

The “Inception” concept comes from the GoogleNet architecture, because GoogleNet is often called Inception-v1, which is introduced by Szegedy et al. [28]. The Inception-v1 model not only reduces the amount of calculations and parameters in a CNN, but at the same time, acquires a top-5 error rate of only 6.67%. Based on the Inception model, Inception-v2, Inception-v3, and Inception-v4 were proposed. In 2015, Szegedy et al. [29] proposed the Inception-v3

Figure 6. General structure of a CNN

architecture, which proposed updates to the Inception module to similarly improve the ImageNet classification accuracy. In 2016, Inception-v4 [30] was introduced. This architecture combined the Inception architecture with residual connections. The Inception-v4 model has many network layers and a complex network structure leading to good image classification performance and easy training for learning. The Inception-v4 module uses a pooling layer with convolution layers stacked together [31]. The convolutions are of varied sizes, that is, 1×1 , 3×3 , and 5×5 . Another salient feature of the Inception module is the use of a bottleneck layer, which uses 1×1 convolutions [29]. The bottleneck layer helps in obtaining a reduction of computational requirements. Additionally, the pooling layer is used for dimension reduction with the module.

2.3.2 Fine-tuning the model

Fine-tuning is a concept used in transfer learning. With the transfer learning, the last few layers of the trained network are removed and retrained with new layers designed as more efficient. Even though the learning process is required in the fine-tuned network, this process is much faster than learning from scratch [32]. Also, fine-tune learning is able to produce more accurate results compared to models trained from scratch.

In this paper, we implement the Inception-v4 model based on the deep learning framework PyTorch [33] and obtain the pre-trained model by training on ImageNet, which has 1,000 classification categories. Then, the last fully connected layer of Inception is fine-tuned, with the number of neurons in the last fully connected layer changed from 1,000 to 5 for Miao embroidery study. Finally, we train the network on Miao embroidery in image set based on the pre-trained model. The Inception-v4 architecture of Miao embroidery image is shown in Fig.7.

2.3.3 Loss function and optimization

(1) Loss function definition

The hypothesis training sample set

$O = \{X, L\}$, $X = \{x^1, x^2, \dots, x^n\}$, represents the feature set of all training samples. $L = \{l^1, l^2, \dots, l^n\}$ indicates the label set of all training samples, and n is the total number of training samples. For any training sample,

its label $l^i \in R^K$; if sample x^i belongs to the j category ($j = 1, 2, \dots, K$), the value of the l^i is 1; otherwise, it is 0. Assuming that the output is $y^i = [y_1^i, y_2^i, \dots, y_K^i]$ of the training sample x^i after the above-mentioned CNN-based seedling embroidery image classification model, the softmax function in equation (1) is here $p^i = [p_1^i, p_2^i, \dots, p_K^i]$. The cross-entropy loss function J can be used to define the above volume-based calculation. The loss function J of the Miao embroidery image classification model of the product neural network is shown in equation (2):

$$p_j^i = \frac{e^{y_j^i}}{\sum_{k=1}^K e^{y_k^i}} \quad (1)$$

$$J = -\sum_{i=1}^n \sum_{j=1}^K l_j^i \log p_j^i \quad (2)$$

(2) Deep model optimization

Before training the proposed CNN model, we first migrate the parameters of the pre-trained model and randomly initialize

Figure 7. The CNN architecture for Miao embroidery image classification

the parameters of the last fully connected layer. Then, a batch sample subset is randomly select from the training sample to train the model. After getting the model output, we use equation (2) to calculate the loss value J and then use stochastic gradient descent and based on equations (3) and (4) to update the model parameters. The pseudo code of the model algorithm is as follows:

$$\frac{\partial J}{\partial W} = \sum_{i=1}^{bs} \sum_{j=1}^K \frac{\partial J}{\partial p_j^i} \cdot \frac{\partial p_j^i}{\partial y_j^i} \cdot \frac{\partial y_j^i}{\partial W} = - \sum_{i=1}^{bs} \sum_{j=1}^K \frac{l_j^i}{p_j^i} \cdot \left(\frac{1}{\sum_{t=1}^K e^{y_t^i}} + e^{y_t^i} \right) \cdot \frac{\partial y_j^i}{\partial W} \quad (3)$$

$$W = W - \partial \cdot \frac{\partial J}{\partial W} \quad (4)$$

Algorithm 1: Miao embroidery image classification algorithm based on CNN optimization

Input: Training sample set $O=\{X,L\}$.

Maximum number of algorithm iterations max_epoch.

Batch value bs.

1: Initialization W , calculate the total number of batches of training samples batch_num= $\lceil n/bs \rceil$.

2: repeat

3: for bn=1 to batch_num do

4: Select bs samples $O_{bz}=\{X_{bz},L_{bz}\}$ from $O=\{X,L\}$ randomly, put X_{bz} in CNN $O=\{X,L\}$ model and calculate the corresponding output Y_{bs} .

5: Calculate P_{bs} by formula (1) and Y_{bs} .

6: Calculate loss function value J by formula (2) and Y_{bs} .

7: Update the parameters of the CNN model W by formulas (3), (4) and stochastic gradient descent.

8: end for

9: until the number of algorithm iterations equals max_epoch

Output: Miao embroidery image classification model W based on CNN.

3. Experiment

3.1 Experimental settings and data split

The experiments are done on the following configuration: A PC with Intel(r)Xeon(r)CPU e3-1050M v5@2.8GHZ, and 32G memory running Windows 10. The development environment is PyTorch, OpenCV [34] is built under Ubuntu 16.2 for image processing, the configuration interface is Python [35].

The Miao embroidery image database is randomly divided in two parts, the training data and the testing data, with a percentage ratio of 80 and 20%, respectively. The training set is used to optimize the deep model, and the test set is used for the performance evaluation of the model. Table 2 shows the number of various types of Miao embroidery used for training and testing in the experiment.

3.2. Image pre-processing

The image pre-processing consist of the data augmentation for CNN models and normalization of the image data [36]. In this experiment, more sample images of Miao embroidery were formed by the data augmentation methods of flipping and cutting. The fully connected layer in the CNN requires a fixed size of the input image, but the sizes of the collected Miao embroidery in the original data set are not completely consistent, so the sizes of some of the images need to be adjusted to meet the CNN model. The size adjustment code is implemented in Python, using the OpenCV library; all the images are zoomed and cropped; and a common image processing "Reshape" operation is used to adjust the image size to 224×224, including reading pictures into the matrix array and adjusting the image size to 299×299×3.

3.3. Evaluation criterion

In the experiment, we use mAP (Mean Average Precision) [37] to evaluate the performance results. mAP is a common metric used for evaluation in the target detection model, with the prediction of position and category of the targets. Here, mAP is employed to evaluate the performance for the classification of Miao embroidery images. The average precision (AP) calculation is formulated in equation (5) if given a query data q :

$$AP(q) = \frac{1}{n_q} \sum_{i=1}^{n_{retrieval}} p_{qi} I(i) \quad (5)$$

Table 2. The details of image dataset used in the experiments

Miao embroidery category	Number of training set	Number of testing set
Wring stitch	364	91
barbola embroidery	244	61
Wrinkling stitch	196	49
Split line stitch	364	91
warp-width counting stitch	196	49

Here, $n_{\text{retrieval}}$ is the total number of images in the retrieved dataset, and n_q is the data number similar to the query data, with p_{q_i} being the number of Top i data that is similar to the query data. $I(i)$ is the indicator function. If the returned data i are similar with q , then $I(i)=1$; otherwise, $I(i)=0$. The corresponding mAP is calculated using equation (6) for query data n_{query} :

$$\text{mAP} = \frac{1}{n_{\text{query}}} \sum_{j=1}^{n_{\text{query}}} \text{AP}(q_j) \quad (6)$$

3.4. Results

In this study, an assessment of the appropriateness of the state-of-the-art deep CNN for the task of Miao embroidery identification is conducted. We fine-tune the Inception-v4 model and compare it with the traditional classification models BoW, FV, and VLAD. To verify the advantages of the fine-tuned CNN for the classification of Miao embroidery, we compare the classification algorithm based on CNN with the algorithm based on BoW, FV, and VLAD. The experiment trained these three models on the training set first, and then used the trained classifier to test the algorithm effect on the test set and calculate the mAP. At the same time, to explore the classification effects of the classification algorithms in different image features dimensions, experiments were conducted under different image feature dimensions to test how mAP changes.

BoW: In the experiment, we first extracted SIFT features of the Miao embroidery images, and then used the k-means method to obtain a visual vocabulary from the feature set. The Miao embroidery image is encoded as a BoW vector, which is used to represent the image. Finally, we use linear SVM to classify the images.

FV: In the experiment, we also firstly extract the local descriptors of SIFT, and then calculate the features' GMM. Then, we use the gradient vector of the likelihood function to calculate and convert the image feature vectors of different lengths to equal lengths. Finally, the normalized feature vector can be used directly in SVM classification.

VLAD: In the experiment, we firstly use k-nearest neighbor (KNN) method to find the nearest clustering center and use the subtract function to calculate the difference between the cluster center of the feature and its nearest neighbor, as well as the add function adds up the difference of the same cluster center. Then, we normalize the matrix into a one-dimensional vector VLAD to do the image classification.

Deep CNN model: In the experiment, we implement the Inception-v4 model and the last fully connected layer of Inception-v4 model is fine-tuned, with the number of neurons in the last fully connected layer changed from 1,000 to 5 for Miao embroidery study. In the training process, the maximum number of iterations of the algorithm is set to 200, and the number of images input to the CNN each time is 128.

3.4.1. The results of the traditional identification model

We calculate and compare the mAP of the BoW, FV, and VLAD for three different image feature dimensions, that is, 500, 1,000, and 2,000, respectively. The differences of the different identification results in the three image feature dimensions based on BoW, FV, and VLAD, are shown in Fig. 8.

As it is illustrated in Fig. 8, when the image feature dimensions are the same, the mAPs of FV are reported significantly higher than those from the BoW and VLAD. It is also seen that when the image feature dimension is 1,000, the three models achieve the highest mAP. BoW, FV, and VLAD are all intermediate representations of images, while FV combines the advantages of both generality and discrimination of the Fisher and GMM architecture. It reflects the frequency of each visual word, as well as encodes the distinctive information of the distribution on the visual words of local features, extracting richer image features than BoW and VLAD. Furthermore, its high-dimensional characteristics make it easier to achieve a good classification result when combining with a simple classifier SVM.

Figure 8. The mAP changes in different image feature dimensions

3.4.2 The comparison of CNN with the traditional identification model

During the training process of deep CNN model, the loss values of the objective function in the training set and the test set are plotted, as well as the accuracy of the test set, as shown in Fig. 9.

The training loss value refers to the error between the predicted class and real result of the training set. Test loss value refers to the error between the test result and real result of the test set, test accuracy refers to the percentage of the randomly selected images for accurate classification that has been made. During the training process, the learning effect is better when the training loss value and the test loss value are smaller, the test accuracy is higher. As shown in Fig. 9(a) and 9(b), the loss values of training and testing are continually decreasing during training. As shown in Fig. 9(c), the accuracy during testing continues to increase until 98.88% eventually.

Figure 9. Variation of testing loss, training loss, and mAP in process of training: (a) is the change of loss value from training set during training process; (b) is the change of loss value from loss value in test dataset; (c) is the mAP of test dataset

To verify the classification performance, we compare the result of the proposed CNN with BoW, FV, and VLAD using image features dimension 1,000. Table 3 illustrates the mAP scores by the four algorithms on the image test set. From Table 3, the mAP of BoW is 79.61%, with VLAD being 81.52%, and FV being 85.77%. The mAP of the Miao image classifier algorithm based on CNN is 98.88%, which more than 13% higher than the accuracy with the other three algorithms. It shows that compared to the traditional classification algorithms, CNNs have sufficient feature extraction capabilities, making it more suitable for the classification of Miao embroidery digital images of southeast Guizhou.

Based on the experiment results, we can see that deep CNN obtained the highest performance for Miao embroidery image recognition. Deep CNN for object recognition is a kind of end-to-end method, in which the input is the image and the output is its predicted category. The learning of feature representation

Table 3. The mAP of different models

Model	mAP
BoW	79.61%
VLAD	81.52%
FV	85.77%
CNN	98.88%

for Miao embroidery image is just a black-box in deep CNNs, and it is known that the low layer of deep CNNs can learn the low-level features (such as texture, color, and edge), the high layer of deep CNNs can learn the high-level features (such as the semantic information). While, the BoW, FV, and VLAD are handcrafted methods. In these methods, the feature representation is designed by researchers for different types of data. However, it is a challenge to design a perfect feature representation for Miao embroidery image. Thus, the high-level features learned by deep CNNs achieve higher performance than vocabulary-based methods for the task of Miao embroidery recognition.

4. Conclusion

In this paper, a CNN-based method is proposed for the classification of different types of Miao embroidery. We firstly established a Miao embroidery image database and manually assigned an accurate category label of Miao embroidery to each image. Then, a pre-trained deep CNN model is fine-tuned based on the established database to learning a more robust deep model to identify the types of Miao embroidery. The experimental results demonstrate that the proposed deep CNN model outperforms the compared three non-deep methods (BoW, FV, and VLAD) and achieved a recognition accuracy of 98.88%. To our best knowledge, this is the first one to apply CNNs on the application of Miao embroidery categories recognition. Moreover, the effectiveness of our proposed method illustrates that the CNN-based approach might be a promising strategy for the discrimination and identification of different other embroidery and national costume patterns.

Acknowledgments

The authors appreciate the helpful comments from the reviewers on improving our work. This work was supported by the Shanghai Design Category IV Peak Discipline Construction Funding Project (No. DC17007), the National Natural Science Foundation of China (61806168), Fundamental Research Funds for the Central Universities (SWU117059), and Venture and Innovation Support Program for Chongqing Overseas Returnees (CX2018075).

References

[1] Liu Bingbing, Yuan Yan. (2019). Try to talk about Historical Memory of Miao Dress Design Symbols. *Guizhou Ethnic Studies*,40:79-82.

[2] Li Ming. (2019). Discission on the Artistic Forms and Causes of Miao Embroidery Crafts in Southeast Guihou. *Folk Art*, 06:90-94.

[3] Zhang Lei, Qin Ziyi. (2019). China National Dyeing and Weaving Intangible Culture Heritage List. *Fashion Guide*, 07(2):13-23.

[4] Yang Shengfeng. (2018). A Brief Analysis of the Development Status of Miao Embroidery Industry in Southeast Guizhou. *Today's Massmedia*, 08:175-176.

- [5] Qi Yuying, Zhang Xiao. (2018). On the Development of Miao Embroidery Industrialization in Southeast Guizhou From the Perspective of Production Protection. *China Collective Economy*, 01:27-29.
- [6] Shang Huifang, Mou Xiaomei, Wang Jiaoyan. (2018). Study on the Present Situation and Countermeasures of Miao Embroidery in Guizhou. *Journal of Qiannan Normal University for Nationalities*, 38(3):122-124.
- [7] TOMOKO Torimaru. One Needle, One Thread. (2011). Guizhou Miao (Hmong) Embroidery and Fabric Piece work from Guizhou, China. China Textile & Apparel Press(Beijing).
- [8] Zhang Y, Jin R, Zhou Z H. (2010). Understanding bag-of-words model: a statistical framework [J].*International Journal of Machine Learning & Cybernetics*, 1(1-4):43-52.
- [9] Gosselin P H, Murray N, Jégou H, et al. (2014). Revisiting the Fisher vector for fine-grained classification. *Pattern Recognition Letters*,49:92-98.
- [10] He Yunfeng, Zhou Ling, Yu Junqing, Xu Tao, Guan Tao.(201). Image Retrieval Based on Locally Features Aggregating [J].*Chinese Journal of Computers*, 34(11):2224-2233.
- [11] Chang L, Deng X M, Zhou M Q, et al. (2016). Convolutional Neural networks in Image understanding. *Acta Automatica Sinica*.
- [12] Artzai P, Maximilian S, Aitor A G, Patrick M, Amaia O B, Jone E. (2019). Crop conditional Convolutional Neural Networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions. *Computers and Electronics in Agriculture*, 167:105093.
- [13] Kattenborn T, Eichel J, Fassnacht F E. (2019). Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. *Scientific Reports*. 9:17656.
- [14] Jia Xiaojun, Ye Lihua, Deng Hongtao, Liu Zihao, Lu Fengjie. (2020). Elements classification of vein patterns using convolutional neural networks for blue calico [J]. *Journal of Textile Research*,41(1).110-117.
- [15] Sheng Jiachuan, Chen Yaqi, Wang Jun, Li Jiang. (2020). Chinese paintings sentiment recognition via CNN optimization with human cognition [J]. *Pattern Recognition and Artificial Intelligence*, 33(2):141-149.
- [16] Li Rongrui, Shi Lin, Zhao Wei. (2019). Minority headdress recognition based on Convolutional Neural Network [J]. *Electronic Sci.&Tech*, 32(2):51-55.
- [17] Ai Hu, Li Fei. (2019). Identification of Guizhou dialect based on improved convolutional neural network [J]. *Modern Information Technology*, 3(1):5-10.
- [18] Wu Huan, Ding Xiaojun, Li Qinman, Du Lei, Zou Fengyuan. (2019). Classification of women's trousers silhouette using convolution neural network CaffeNet model [J]. *Journal of Textile Research*, 40(4):117-121.
- [19] Wang Fei, Jin Xiangyu. (2017). Identification of cashmere and wool based on convolutional neuron networks and deep learning theory [J]. *Journal of Textile Research*, 38(12):150-156.
- [20] Liu Zhengdong, Liu yihan, Wang Shouren. (2019). Depth learning method for suit detection in images [J]. *Journal of Textile Research*, 40(04):158-164.
- [21] He Xiaoyun, Wei Ping, Zhang Lin, Deng Binyou, Pan Yunfeng, Su Zhenwei. (2018). Detection method of foreign fibers in seed cotton based on deep-learning. *Journal of Textile Research*, 39(6):131-135.
- [22] Wang Wenwen, Gao Chang, Liu Jihong. (2018). Position recognition of spinning yarn breakage based on convolution neural network. *Journal of Textile Research*, 39(6):136-141.
- [23] Wang Shanna, Zhang Huaxiong, Kang Feng. (2018). Emotion classification of necktie pattern based on convolution neural network. *Journal of Textile Research*, 39(8):117-123.
- [24] Corona, E; Alenya, G; Gabas, A; Torras, C. (2018). Active garment recognition and target grasping point detection using deep learning. *Pattern Recognition*, 74(2):629-641.
- [25] Zhao Yudi, Hao Kuangrong, He Haibo, Tang Xuesong, Bei Bing. (2020). A visual long-short-term memory based integrated CNN model for fabric defect image classification. *Neurocomputing*, (380)03:259-270.
- [26] Jing Jun-Feng, Ma Hao, Zhang Huanhuan. (2019). Automatic fabric defect detection using a deep convolutional neural network. *Coloration Technology*, 135(6):213-223.
- [27] Szegedy C, Ioffe S, Vanhoucke V, et al. (2017). Inception-v4, inception-ResNet and the impact of residual connections on learning .31st AAAI Conference on Artificial Intelligence, AAAI 2017:4278-4284.
- [28] Szegedy, C., Com, S.G.(2017). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv*,37.
- [29] Szegedy, C., Vanhoucke, V., Shlens, J. (2016). Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*,2818-2826.\
- [30] Szegedy, C., Iofffe, S., Vanhoucke, V., Alemi, A.A.(2017). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In: AAAI, 4:12.
- [31] Kuang H L, Liu C R, Chan L L, Yan H. (2018). Multi-class fruit detection based on image region selection and improved object proposals. *Neurocomputing*. 283:241-255.
- [32] Mohanty, S.P., Hughes, D.P., Salathé, M. (2016). Using deep learning for image-based plant disease detection [J]. *Frontiers in Plant Science*, 7 (9): 1–7.
- [33] <http://pytorch.org/>.
- [34] <http://opencv.org/>.
- [35] <http://python.org/>.
- [36] Li Shengwang, Han Qian. (2018). Image Processing Technology Based on Deep Learning. *Digital Technology &Application*,36(09):5-66.
- [37] Cheng Deng, Zhaojia Chen, Xianglong Liu, Xinbo Gao, Dacheng Tao. (2018). Triplet-Based Deep Hashing Network for Cross-Modal Retrieval. *IEEE Transactions on Image Processing*, 27 (8), 3893 - 3903.