Research on the Clothing Classification of the She Ethnic Group in Different Regions Based on FPA-CNN

Xiaojun Ding 1,2,*, Tao Li 1,2, Jingyu Chen 1,3, Ling Ma 1,2 and Fengyuan Zou 1,2,3

1 School of Fashion Design and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China
2 Engineering Research Center of Clothing of Zhejiang Province, Zhejiang Sci-Tech University, Hangzhou 310018, China
3 Key Laboratory of Silk Culture Inheriting and Products Design Digital Technology, Ministry of Culture and Tourism, Zhejiang Sci-Tech University, Hangzhou 310018, China
* Correspondence: dxj_tina@126.com

Abstract: In order to achieve the effective computer recognition of the She ethnic clothing from different regions through the extraction of color features, this paper proposes a She ethnic clothing classification method based on the Flower Pollination Algorithm-optimized color feature fusion and Convolutional Neural Network (FPA-CNN). The method consists of three main steps: color feature fusion, FPA optimization, and CNN classification. In the first step, a color histogram and color moment features, which can represent regional differences in She ethnic clothing, are extracted. Subsequently, FPA is used to perform optimal weight fusion, obtaining an optimized ratio. Kernel principal component analysis is then applied to reduce the dimensionality of the fused features, and a CNN is constructed to classify the She ethnic clothing from different regions based on the reduced fused features. The results show that the FPA-CNN method can effectively classify the She ethnic clothing from different regions, achieving an average classification accuracy of 98.38%. Compared to SVM, BP, RNN, and RBF models, the proposed method improves the accuracy by 11.49%, 7.7%, 6.49%, and 3.92%, respectively. This research provides a reference and guidance for the effective recognition of clothing through the extraction of color features.

Keywords: She ethnic clothing; color feature; Flower Pollination Algorithm; Convolutional Neural Network; classification accuracy

1. Introduction

The She ethnic group is a distinct minority in southeastern China, mainly distributed in five regions: Jingning, Fu’an, Luoyuan, Xiapu, and Fuding [1]. She ethnic clothing has a continuous lineage, but in different regions, it exhibits different characteristics [2]. Accordingly, it is of great significance to distinguish the clothing features and differences of the She ethnic group in different regions for the protection, inheritance, design innovation, and sustainable development of She ethnic clothing [3–5].

In the past, the classification of She ethnic clothing relied mainly on manual experience, which was not only time-consuming and laborious, but also highly dependent on the subjective judgment and experience of the people making the classification. With the development of computer vision technology, researchers have achieved certain research results in the objective recognition of traditional clothing using feature extraction and clothing recognition techniques.

The objective of the feature extraction always includes shape, texture, and color. Color, as one of the most important criteria for ethnic clothing recognition, can reflect the personalized information of the ethnic group well, since color features have advantages such as good stability, scale invariance, and rotational invariance in comparison with the other two features [6]. Therefore, some researchers have proposed that it is more accurate for classification results to use color features as the main criterion for clothing recognition.
recognition. Xing et al. [7] used the mean-shift clustering algorithm to extract the color features of traditional clothing, achieving the color expression of traditional clothing. Liu et al. [8] described the color of fabric images using color histogram features and successfully extracted the dominant color series using a three-level weight method. Ding et al. [9] transformed color images into grayscale images and achieved the recognition of She ethnic clothing by extracting SIFT features combined with a bag-of-words approach. Zhang et al. [10] proposed an identification system based on the fusion of color-dominant series and color moment features, which achieved high accuracy using color features. Huang et al. [11] extracted the fused bundled features that represent image position and color features, while Gupta et al. [12] integrated clothing color and texture features and obtained good retrieval results using computational vector distance. Although the above studies achieved a feature fusion to some extent, the proportional combination of single features mainly relies on subjective judgment. Considering the limitation of effectively determining the weight combination, this study introduces the Flower Pollination Algorithm (FPA) to automatically optimize the parameters instead of manual intervention. Compared to other optimization algorithms, the FPA has advantages such as fewer parameters, good stability, and resistance to premature convergence, making it suitable for global and local optimization [13].

In terms of technical recognition, Nawaz et al. [14] proposed a method using Convolutional Neural Networks (CNN) based on GoogleNet’s inception for the automatic classification of traditional ethnic clothing images in Bangladesh, achieving recognition of upper garments, lower garments, and suits. Huo et al. [15] used part detection and feature fusion techniques to classify the images of 11 representative ethnic clothing styles, including Miao, Mongolian, and Korean ethnic groups. Sun et al. [16] used a multi-task neural network to extract clothing attribute features and achieved recognition for 14 ethnic groups, including Bai, Gino, and Tibetan ethnic groups. Wu et al. [17] used semantic attributes and multi-task learning to recognize 25 ethnic minority clothing groups in Yunnan, and the recognition accuracy reached 82.5~88.4%. Currently, research on clothing recognition mainly focuses on different types of clothing styles [18–20], such as upper garments and lower garments or clothing styles with significant differences in contours [21–23]. Recognizing similar styles with small differences in the same category poses a challenge. Deep learning plays a key role in computer recognition. Common classifiers include recurrent neural networks (RNNs), Convolutional Neural Networks (CNNs), and artificial neural networks (ANNs). Compared to RNNs and ANNs, CNNs have better generalization [24,25], which provides a reference for this study.

She ethnic clothing from different regions has similarities in shape and texture features, but their color features have a certain degree of discriminability. The innovation of this work lies in proposing a recognition method for She ethnic clothing based on FPA-optimized color feature fusion and a CNN. The main contributions are as follows: (1) A method for color feature fusion and dimension reduction was proposed. The extracted color histogram features and color moment features were normalized and then optimized using the Firefly Algorithm (FPA) to obtain optimal weights. To reduce the dimensionality of the input algorithm, the proposed method also employed kernel principal component analysis (PCA) to fuse high-dimensional features, thereby reducing the complexity of the system. (2) When classifying clothing from different regions of the She ethnic group, the use of receptive fields and weight-sharing techniques reduced the number of required weights and the computational burden during network training. Through supervised learning, the information retrieval capability of the classification model was further enhanced, effectively improving the classification accuracy.

To do this, we structure the rest of this paper as follows: In Section 2, we briefly summarize the research framework, and then the experimental methods (color feature extraction, fusion, dimensionality reduction, etc.) are elaborated. In Section 3, the color analysis and She ethnic clothing classification and comparison are conducted. In Section 4,
we discuss the experimental results, limitations, and future works. Finally, we present some conclusions in Section 5.

2. Experiments

2.1. Experimental Procedure

In this study, we propose a She ethnic clothing recognition method based on FPA-optimized color feature fusion and a CNN to address the challenge of effectively classifying traditional clothing from different regions with similar shape and texture features. The experimental procedure consists of three major steps: color feature extraction, FPA optimization and fusion, and CNN classification, as shown in Figure 1.

Figure 1. Overall proposed method.

2.2. Experimental Samples

A total of 370 She ethnic clothing images were selected for the experiment, including 74 samples from each of the five regions in China: Jingning, Fu’an, Luoyuan, Xiapu, and Fuding. These images were obtained from the China She Ethnic Museum, relevant monographs [26,27], and other sources. The dataset consists of images of female upper garments of the She ethnic group. The samples were divided into a training set (80%) and a test set (20%). Figure 2 shows five typical She ethnic clothing images.

The experimental software platform used was Matlab R2022a running on a Win64 system. The hardware configuration included an Intel i5 processor and 8 GB of RAM. These resources were utilized for the implementation of the experimental procedure and analysis in the subsequent sections.
R (1) Color Space Quantization

Compared to the RGB and CMYK color space models, the HSV color space, which consists of three color channels: Hue, Saturation, and Value, is more consistent with human visual perception [28–30]. This separation through H, S, and V enables a better capturing of various colors in clothing, especially in the case of the vibrant and rich colors frequently used in ethnic clothing [31]. Therefore, before the extraction of color features, the images are converted from the RGB color space to the HSV color space using the conversion formulas shown in Equations (1)–(3):

\[
H = \begin{cases} 
\frac{G-B}{\max(R,G,B)-\min(R,G,B)} & \text{if } \max(R,G,B) = \min(R,G,B) \\
\frac{60^\circ \times \frac{G-B}{\max(R,G,B)-\min(R,G,B)}}{\max(R,G,B)} & \text{if } \max = R \text{ and } G \geq B \\
\frac{60^\circ \times \frac{B-R}{\max(R,G,B)-\min(R,G,B)}}{\max(R,G,B)} + 360^\circ & \text{if } \max = R \text{ and } G < B \\
\frac{60^\circ \times \frac{B-R}{\max(R,G,B)-\min(R,G,B)}}{\max(R,G,B)} + 120^\circ & \text{if } \max = G \\
\frac{60^\circ \times \frac{G-B}{\max(R,G,B)-\min(R,G,B)}}{\max(R,G,B)} + 240^\circ & \text{if } \max = B
\end{cases}
\]

\[
S = \begin{cases} 
0, & \text{if } \max = 0 \\
1 - \frac{\min(R,G,B)}{\max(R,G,B)}, & \text{if } \max \neq 0
\end{cases}
\]

\[
V = \frac{\max(R,G,B)}{255}
\]

where R, G, and B represent the red, green, and blue channel values, respectively, “max” and “min” represent the maximum and minimum values among R, G, and B. H belongs to the range \([0^\circ, 360^\circ]\), while S and V belong to the range [0, 1].

2.3. Color Space Quantization

In order to describe the clothing information of the She ethnic group in different regions with higher accuracy, this paper proposes a method that combines color histograms and color moment features based on HSV color space to extract the color features of the She ethnic group clothing in different regions.

2.3.1. Color Feature Extraction of She Ethnic Groups in Different Regions

The distribution of pixels in the image’s H, S, and V color channels can be described using a color histogram. In order to obtain effective features of the She ethnic clothing images, this study employs a non-uniform quantization method to process the H, S, and V channels, reducing the excessive dimensionality of the feature vector and improving the

![Figure 2. Samples for the experiment. (a) Jingning; (b) Fu’an; (c) Luoyuan; (d) Xiapu; (e) Fuding.](image-url)
efficiency of classifier construction and recognition accuracy. In the experiment, H, S, and V were non-uniformly quantized into 16, 4, and 4 channels, respectively [32]. The three color channels were converted into a one-dimensional color feature vector, as shown in Equation (4), to further reduce the feature dimensionality:

\[ L = 16H + 4S + V \]  

(4)

where \( L \in [0, 1, \ldots, 255] \) means that the quantized color space consists of 256 color feature values, and the color histogram feature vector is obtained by \( \{L_0, L_1, \ldots, L_{255}\} \). The extracted histogram features for different regions of She ethnic clothing are shown in Figure 3.

![Color histograms](a), (b), (c), (d), (e)

Figure 3. Color histograms. (a) Jingning; (b) Fu’an; (c) Luoyuan; (d) Xiapu; (e) Fuding.

(3) Color moment features

Color moments are introduced as an effective supplement to compensate for the problem of ignoring pixel position distribution in color histograms. Color information is mainly distributed in low-order moments [33]. Therefore, in the HSV color space model, the first-order moment \( \mu_i \), second-order moment \( \sigma_i \), and third-order moment \( \omega_i \) are extracted from the image. The calculation formulas are shown in Equations (5)–(7). Since each pixel has three color components in HSV (Hue, Saturation, and Value), there are nine feature vectors for color moments \( \{\mu_H, \sigma_H, \omega_H, \mu_S, \sigma_S, \omega_S, \mu_V, \sigma_V, \omega_V\} \):

\[ \mu_i = \frac{1}{N} \sum_{j=1}^{N} P_{ij} \]  

(5)

\[ \sigma_i = \left[ \frac{1}{N} \sum_{j=1}^{N} (P_{ij} - \mu_i)^2 \right]^{1/2} \]  

(6)
Therefore, kernel principal component analysis is employed in this study to reduce the number of pixels obtained from Equation (4), \( t \) represents the value with the historical best value until the maximum number of iterations is reached. The position update is performed by comparing the current fitness value to the historical best fitness value. For the fitness function, the optimal parameter combination is obtained, and the FPA model is established for feature fusion. The optimal parameters are then used in the CNN model to recognize and classify the She ethnic clothing.

To obtain a comprehensive representation of She ethnic clothing images, the color histograms and color moment features are normalized and fused using the FPA to optimize the fusion weights \[34\]. The weighted fusion of color features contains more information \[35,36\]. The fused feature is denoted as \( \text{cnn} \) and can be expressed as follows in Equation (8):

\[
f = \text{accuracy}\{\text{cnn}(a \ast t_1 + \beta \ast t_2)\}
\]

where \( f \) represents the objective function, \( t_1 \) represents the color histogram feature \( (L) \) values obtained from Equation (4), \( t_2 \) represents the color moment feature \( \{\mu_H, \sigma_H, \omega_H, \mu_S, \sigma_S, \omega_S, \mu_V, \sigma_V, \omega_V\} \), \( a \) and \( \beta \) are the weights, \( \text{cnn} \) is the classifier, and accuracy represents the evaluation metric for the classification results.

The FPA algorithm is utilized for optimization by updating the parameters through the fitness function. The position update is performed by comparing the current fitness value with the historical best value until the maximum number of iterations is reached. The optimal parameter combination is obtained, and the FPA model is established for feature fusion. The optimal parameters are then used in the CNN model to recognize and classify the She ethnic clothing.

(2) Feature Dimensionality Reduction

The color features result in a 256-dimensional color histogram feature vector and a 9-dimensional color moment feature vector. Such high-dimensional data may lead to long training times and impact the recognition accuracy of the classification model. Therefore, kernel principal component analysis is employed in this study to reduce the dimensionality of the She ethnic clothing color feature vectors. This technique yields a set of eigenvalues and eigenvectors, which can be used to project the data into a lower-dimensional space while preserving the most important information. The dimensionality of the feature vector is reduced by selecting the top \( k \) eigenvectors with the highest eigenvalues, where \( k \) is a parameter that can be determined based on the desired trade-off between dimensionality reduction and information loss. The resulting lower-dimensional feature vector can then be used for classification tasks, leading to improved performance and reduced computational requirements.
of low-dimensional, independent components that can explain most of the variance in the original data, effectively achieving high-dimensional data reduction [37].

2.3.3. The Construction of the She Ethnic Clothing Classifier Based on the CNN

Considering that the She ethnic clothing images consist of repetitive local information with low complexity, traditional approaches based on shape and highly complex features are not suitable. Instead, a more in-depth local feature extraction is required. The essence of CNN image classification lies in the convolutional layers that extract local features from images and pass them to fully connected layers to obtain global features of the clothing image, thereby achieving the fine-grained classification of the clothing images. Convolutional Neural Networks (CNNs) are a type of deep neural network with a convolutional operation, and they have a deep structure [38]. In this network, the convolutional structure is used to reduce the number of network parameters and the memory usage of deeper layers. The CNN mainly consists of the following two layers.

(1) Convolutional Layer

The convolutional layer is a hidden layer that performs convolutional operations and serves as the fundamental structure of the CNN. The function of the convolutional layer is to connect each neuron in the current layer with the neurons in the previous layer within a convolutional window. It adopts parameter sharing to allocate the same set of parameters to neurons in the same layer [39]. The convolutional process can be expressed as follows in Equation (9):

$$h^k_{ij} = f \left( \sum_{i=1}^{k} W^k * x_{ij} + b_k \right)$$

where $h^k_{ij}$ represents the element at the $k$-th layer, the $i$-th row, and $j$-th column, $W^k$ represents the value of the input data at position $(i,j)$, $f$ denotes the activation function, and $b_k$ represents the bias of the convolutional kernel. The convolutional layer has two convolutional kernels and two channels, with each channel corresponding to a convolutional kernel. The activation functions used in this study are ReLU and sigmoid functions.

(2) Pooling Layer

The pooling layer exhibits the property of local connections. Its primary function is to perform down-sampling on the data, achieve feature selection, and further prevent overfitting [40]. The pooling layer differs from the convolutional layer in that it does not have weighted coefficients, and therefore it does not require training to optimize the computation results. The general expression of the commonly used pooling method, denoted as $L_p$, is as follows:

$$A^k_l(i, j) = \left[ \sum_{x=1}^{f} \sum_{y=1}^{f} A^k_l(s_0i + x, s_0j + y) \right]^{1/p}$$

where $s_0$ represents the parameters of the pooling layer and $p$ is a predetermined parameter. When $p \to \infty$, $L_p$ is referred to as max pooling, which takes the maximum value within the pooling region. When $p = 1$, $L_p$ calculates the average value within the pooling region and is known as average pooling.

Figure 5 shows the CNN network structure, including the five-layer convolution layer and the three-layer maximum pooling layers. Each convolutional kernel has a size of $3 \times 3$. The level uses ReLU as the activation function to construct a maximum pooling layer to reduce the output dimension. The convolution step size of the max pooling layer is set to 2.
3. Results and Discussion

3.1. Color Analysis of She Ethnic Clothing

To ensure the completeness of color feature extraction for She ethnic clothing, color analysis was conducted in the HSV color space using color quantization. The color distribution of the She ethnic clothing in different regions is shown in Table 1.

Table 1. Color analysis of the She ethnic clothing in different regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Color</th>
<th>ΔE</th>
<th>Rate/%</th>
<th>H/S/V</th>
<th>Color</th>
<th>ΔE</th>
<th>Rate/%</th>
<th>H/S/V</th>
<th>Color</th>
<th>ΔE</th>
<th>Rate/%</th>
<th>H/S/V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jingning</td>
<td>black</td>
<td>1.5</td>
<td>96.65</td>
<td>[207, 22, 6]</td>
<td>red</td>
<td>2.9</td>
<td>1.20</td>
<td>[17, 38, 81]</td>
<td>brown</td>
<td>2.3</td>
<td>1.10</td>
<td>[6, 51, 51]</td>
</tr>
<tr>
<td>Fu’an</td>
<td>black</td>
<td>1.7</td>
<td>96.43</td>
<td>[45, 49, 5]</td>
<td>red</td>
<td>2.3</td>
<td>1.36</td>
<td>[0, 71, 66]</td>
<td>brown</td>
<td>4.5</td>
<td>1.31</td>
<td>[7, 45, 61]</td>
</tr>
<tr>
<td>Luoyuan</td>
<td>black</td>
<td>0.5</td>
<td>55.91</td>
<td>[239, 46, 1]</td>
<td>red</td>
<td>3.9</td>
<td>10.76</td>
<td>[0, 50, 88]</td>
<td>brown</td>
<td>3.4</td>
<td>8.79</td>
<td>[358, 35, 91]</td>
</tr>
<tr>
<td>Xiapu</td>
<td>black</td>
<td>1.3</td>
<td>5.41</td>
<td>[21, 13, 67]</td>
<td>red</td>
<td>3.2</td>
<td>4.59</td>
<td>[55, 44, 84]</td>
<td>brown</td>
<td>2.5</td>
<td>4.16</td>
<td>[9, 75, 78]</td>
</tr>
<tr>
<td>Fuding</td>
<td>black</td>
<td>2.0</td>
<td>94.23</td>
<td>[69, 66, 3]</td>
<td>red</td>
<td>1.8</td>
<td>2.31</td>
<td>[8, 66, 56]</td>
<td>brown</td>
<td>2.9</td>
<td>1.82</td>
<td>[16, 49, 57]</td>
</tr>
<tr>
<td></td>
<td>red</td>
<td>1.1</td>
<td>0.81</td>
<td>[12, 57, 41]</td>
<td>red</td>
<td>2.5</td>
<td>4.13</td>
<td>[2, 79, 72]</td>
<td>brown</td>
<td>1.2</td>
<td>3.24</td>
<td>[0, 84, 53]</td>
</tr>
<tr>
<td></td>
<td>brown</td>
<td>0.6</td>
<td>87.66</td>
<td>[3, 71, 1]</td>
<td>red</td>
<td>3.4</td>
<td>1.76</td>
<td>[358, 70, 32]</td>
<td>brown</td>
<td>3.1</td>
<td>1.31</td>
<td>[5, 56, 62]</td>
</tr>
</tbody>
</table>

From Table 1, it can be seen that the color, proportion, color sequence, and color matching from different regions of She nationality are quite different. Therefore, different She ethnic clothing has its own original color information. It can also be observed that the main colors of She ethnic clothing are black and dark blue, with relatively low values of lightness. The proportions of the main colors range from 55.91% to 96.68%. The proportion of the main color in the Luoyuan region is significantly lower than that in other regions, while the Jingning region has the highest proportion, reaching 96.65%. The complementary colors of She ethnic clothing exhibit a rich and colorful palette, mainly consisting of traditional Chinese red shades (such as cheek red, crimson, and vermilion) and traditional brown shades (such as ocher, tea color, and tortoiseshell). The complementary colors form a distinct contrast in saturation and lightness with the main colors, and their proportions range from 3.35% to 44.09%. Among them, the Luo Yuan region has the richest matching colors. On the one hand, from the perspective of ethnic culture, She ethnic clothing has a...
strong ethnic style and ethnic characteristics. On the other hand, the She ethnic clothing in different regions also shows diversity, with variations in proportion and color schemes, which is consistent with the conclusions of [41,42]. This provides a reliable basis for the subsequent establishment of a classification model.

3.2. She Ethnic Clothing Classification Based on FPA-CNN

3.2.1. Classification Index Selection

The CNN consists of five convolutional layers and three max pooling layers. Each convolutional kernel has a size of $3 \times 3$. The ReLU activation function is used in the layers, and max pooling layers are constructed to reduce the output dimensions. The convolutional stride in the max pooling layers is set to 2. The evaluation metric for the experiment is image classification accuracy, which is calculated using the following Equation (11):

$$\text{Accuracy} = \frac{\sum(Y_{\text{label}} = y_{\text{label}})}{\text{Length}(y_{\text{label}})} \times 100\%$$

where $Y_{\text{label}}$ represents the correctly classified samples and $y_{\text{label}}$ represents the samples in the test set.

3.2.2. Analysis of the She Ethnic Clothing Classification Results

Table 2 presents the classification results for the She ethnic clothing in different regions. From Table 2, it can be observed that the FPA-CNN method developed in this study achieves good recognition results for She ethnic clothing in the five regions. Among the 74 samples of She ethnic clothing, 73 are correctly classified, resulting in an overall classification accuracy of 98.65%. However, there is a case of misclassification. For instance, one item of She ethnic clothing from the Xiapu region is incorrectly classified as Fuding clothing.

Table 2. Clothing classification of the different branches of the She ethnic group.

<table>
<thead>
<tr>
<th>Category</th>
<th>Jingning</th>
<th>Fu’an</th>
<th>Luoyuan</th>
<th>Xiapu</th>
<th>Fuding</th>
<th>Classification Accuracy/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jingning</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Fu’an</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Luoyuan</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Xiapu</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>1</td>
<td>93.75</td>
</tr>
<tr>
<td>Fuding</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>100</td>
</tr>
</tbody>
</table>

There are two main reasons for the misclassification. Firstly, compared to Xia Pu clothing, Fuding clothing has two “Red bayberry Flower” features on the collar. Xiapu clothing does not have the “Red bayberry Flower,” and the proportion of this color feature is relatively small compared to the entire item of clothing. This makes it difficult to establish effective feature differences, leading to difficulties in distinguishing clothing from these two regions and resulting in incorrect classification. Secondly, the collection of She ethnic clothing images is incomplete, which hampers the ability to fully capture the distinct features of clothing from different regions. This incomplete representation during the image collection and segmentation processes leads to feature omissions and incomplete transformation into image information. As a result, there may be similarities between the color histograms and color moment features of Fuding and Xiapu clothing, leading to the misclassification.

In summary, while the FPA-CNN method demonstrates high accuracy in classifying She ethnic clothing from different regions, there are still challenges in distinguishing certain clothing items due to subtle feature differences and incomplete image representation.
3.3. Comparative Analysis of She Ethnic Clothing Classification

3.3.1. Comparative Analysis of Parameter Optimization Methods

To address the subjectivity issue in manually determining parameter combinations, intelligent optimization algorithms are integrated into the parameter optimization process, considering the specific application scenarios. Common intelligent optimization algorithms include the Flower Pollination Algorithm (FPA) \[43,44\] and Particle Swarm Optimization (PSO) \[45\], the Bat algorithm (BA) \[46\], Harmony Search (HS) \[47\], and the Artificial Bee Colony (ABC) \[48\]. Compared to these algorithms, the FPA has advantages such as fewer parameters, better stability, and less susceptibility to premature convergence. The FPA also exhibits excellent global and local search capabilities, achieving high optimization efficiency \[49\]. The QFPA is an improved algorithm based on the FPA by Ebubekir Kaya et al. in 2022 \[50\].

To validate the effectiveness of the FPA, a comparison experiment was conducted on the minimum values of the second-order function using the FPA and QFPA, PSO, BA, HS, and ABC algorithms, respectively, as shown in Equation (12).

\[
F = -20 \exp \left( -0.2 \sqrt{\sum_{i=1}^{2} x_i^2} \right) - \exp \left( \frac{1}{2} \sum_{i=1}^{2} \cos(2\pi x_i) \right) + e + 20 \tag{12}
\]

where \(F\) represents the objective function, which is a second-order function, \(x_i\) represents the independent variables, specifically \(x_1\) and \(x_2\), and \(e\) represents the natural constant.

Algorithm 1 presents the pseudo code of the FPA; the basis of the FPA relies on both local and global pollination mechanisms. Global pollination occurs in larger areas through biotic factors; abiotic factors enable the occurrence of local pollination in more limited areas.

```
Algorithm 1 Pseudo code of FPA [13]
Objective min or max f(X), X = (x_1, x_2, \ldots, x_d)
Initialize a population of n flowers/pollen gametes with random solutions
Find the best solution \(g^*\) in the initial population
Define a switch probability \(p \in [0, 1]\)
While \((t < \text{MaxGeneration})\)
  for \(i = 1:n\) (all \(n\) flowers in the population)
    if rand < \(p\),
      Draw a (d-dimensional) step vector \(L\) which obeys a Lévy distribution
      Global pollination via \(X_{t+1}^i = X_t^i + L(g^* - X_t^i)\)
    else
      Draw \(\epsilon\) from a uniform distribution in \([0, 1]\)
      Randomly choose \(j\) and \(k\) among all the solutions
      Do local pollination via \(X_{t+1}^i = X_t^i + \epsilon(X_t^j - X_t^k)\)
    end if
    Evaluate new solutions
    If new solutions are better, update them in the population
  end for
  Find the current best solution \(g^*\)
end while
```

Figure 6a shows the optimization process of the FPA, QFPA, PSO, BA, HS, and ABC algorithms. Figure 6b presents the optimization result obtained by the FPA. As shown in Figure 5a, compared with the PSO, BA, HS, and ABC algorithms, the FPA and QFPA have faster optimization speeds. Compared with the QFPA, the FPA has a slower optimization speed, but its optimization results are closer to the global optimal. After 38 iterations, the FPA finds the minimum value of 0 and maintains it throughout the remaining iterations. The optimal solution obtained with the FPA is highly accurate, with coordinates precise
up to nine decimal places, indicating its ability to approximate the global minimum more effectively. Therefore, this paper adopts the FPA for She ethnic clothing color feature fusion.

![Graph showing optimization processes for FPA, QFPA, PSO, BA, HS, and ABC algorithms.](image)

**Figure 6.** Optimization process and results. (a) Comparison of optimization processes for FPA, QFPA, PSO, BA, HS, and ABC algorithms; (b) FPA optimization results.

### 3.3.2. Comparative Analysis of Feature Fusion Methods

To validate the effectiveness of the FPA color feature fusion method, this study explores the combination of color histograms and color moments by setting multiple sets of weights. The results are then compared with the results obtained from the FPA color feature fusion approach. The fused features from each set are subjected to kernel principal component analysis to construct the input set for the CNN classification model. The cumulative contribution rate of the features is approximately 95%, ensuring the objectivity and consistency of the experiments.

The testing samples were evaluated in terms of training time and classification accuracy. The results are summarized in Table 3.

<table>
<thead>
<tr>
<th>Color Feature</th>
<th>Method</th>
<th>Training Time/s</th>
<th>Classification Accuracy/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-color feature</td>
<td>$t_1$</td>
<td>24.56</td>
<td>86.49</td>
</tr>
<tr>
<td></td>
<td>$t_2$</td>
<td>22.37</td>
<td>81.08</td>
</tr>
<tr>
<td></td>
<td>$t_1 + t_2$</td>
<td>23.45</td>
<td>87.83</td>
</tr>
<tr>
<td></td>
<td>$0.9 \times t_1 + 0.1 \times t_2$</td>
<td>21.69</td>
<td>89.19</td>
</tr>
<tr>
<td></td>
<td>$0.8 \times t_1 + 0.2 \times t_2$</td>
<td>21.42</td>
<td>90.54</td>
</tr>
<tr>
<td></td>
<td>$0.7 \times t_1 + 0.3 \times t_2$</td>
<td>21.39</td>
<td>93.24</td>
</tr>
<tr>
<td></td>
<td>$0.6 \times t_1 + 0.4 \times t_2$</td>
<td>22.07</td>
<td>94.59</td>
</tr>
<tr>
<td></td>
<td>$0.5 \times t_1 + 0.5 \times t_2$</td>
<td>21.59</td>
<td>91.89</td>
</tr>
<tr>
<td></td>
<td>$0.4 \times t_1 + 0.6 \times t_2$</td>
<td>22.17</td>
<td>90.54</td>
</tr>
<tr>
<td></td>
<td>$0.3 \times t_1 + 0.7 \times t_2$</td>
<td>22.49</td>
<td>86.48</td>
</tr>
<tr>
<td></td>
<td>$0.2 \times t_1 + 0.8 \times t_2$</td>
<td>21.65</td>
<td>83.78</td>
</tr>
<tr>
<td></td>
<td>$0.1 \times t_1 + 0.9 \times t_2$</td>
<td>21.89</td>
<td>82.43</td>
</tr>
<tr>
<td>Multi-color feature fusion</td>
<td>Proposed method</td>
<td>13.25</td>
<td>97.29</td>
</tr>
</tbody>
</table>

**Table 3.** Comparison of classification effects of different features.
According to Table 3, it can be observed that the classification accuracy of the color histogram feature is higher than that of the color moment feature when used individually. This indicates a positive correlation between the richness of information and classification accuracy. However, due to the high dimensionality of the color histogram feature, the computational complexity is high, resulting in a longer computation time, even after dimensionality reduction through the kernel principal component analysis.

In comparison with using a single feature, the fusion of multiple color features, achieved by weighting the color histograms and color moment features, effectively improves the classification accuracy for She clothing. The experiments also demonstrate that increasing the weight of the color moment feature can enhance the classification accuracy. However, when $t_2$ reaches 0.5, the classification accuracy begins to gradually decrease. This is because, as the weight of the color histogram feature decreases, the feature information contained in the fusion of multiple color features is reduced, leading to a decrease in classification accuracy.

In this study, the FPA was used to optimize the best weight fusion, with a focus on adopting an adaptive mechanism for optimizing the color feature weights. The experimental results show that when the weights of the color histograms and color moment features are 0.645 and 0.355, respectively, the FPA color feature fusion achieves the highest classification accuracy of 97.29%, with a training time of only 13.25 s. This is because FPA color feature fusion can extract data features and obtain optimal weights for feature fusion more efficiently. Therefore, the effectiveness of the proposed FPA color feature fusion method is demonstrated, and it is adopted as the input feature for the She ethnic clothing classifier in this study.

### 3.3.3. Comparative Analysis of Classification Models

In order to validate the effectiveness of the classification models, this study conducted 10 classification experiments on CNN, SVM, BP, RNN, and RBF models. The test accuracy and average values obtained from these experiments are presented in Figure 7.

![Classification accuracy of different models](image)

**Figure 7.** Classification accuracy of different models.

From Figure 7, it can be observed that, in the 10 experiments, the average classification accuracy of SVM is 86.89%. This is because the SVM model uses non-linear planes for class separation, which leads to unsatisfactory classification results for She ethnic clothing classification. Due to the presence of a large amount of redundant information in the
training process, it is necessary to optimize the classifier performance by introducing deep learning networks [51]. Compared to the SVM, the classification accuracy of BP, RNN, and RBF neural network models were significantly improved, with average classification accuracy increased by 3.79%, 5%, and 7.57%, respectively. This indicates that neural networks have higher applicability in She ethnic clothing classification, especially in the fusion of color histograms and color moment features, where they exhibit better learning performance. However, the highest average classification accuracy achieved by the three aforementioned neural network models is only 94.46%.

The FPA-CNN model proposed in this paper consistently achieved the highest classification accuracy in all 10 tests, reaching a maximum of 98.65%. The average classification accuracy also reached 98.38%, showing improvements of 11.49%, 7.7%, 6.49%, and 3.92% compared to the corresponding results from the SVM, BP, RNN, and RBF models, respectively. The effectiveness of the proposed method in the She ethnic clothing classification is significant. This is because, compared to traditional classification algorithms, the proposed method employs receptive fields and weight-sharing techniques, reducing the number of weights required for training the network and the associated burden. Additionally, through supervised learning, the proposed method further enhances the information acquisition capabilities of the classification model, as a result of which the classification accuracy is effectively improved.

To visually validate the effectiveness of the proposed method in She ethnic clothing feature classification, this study performed a kernel principal component analysis on the class features from the CNN, SVM, BP, RNN, and RBF classification results and simplified them into a two-dimensional plane. Based on this, comparative experiments were conducted using different classifiers, and the results are shown in Figure 8.

![Figure 8](image_url)

**Figure 8.** Comparison of different classification methods and visualization. (a) SVM; (b) BP; (c) RNN; (d) RBF; (e) CNN.
From Figure 8, we can see that the SVM, BP, RNN, and RBF all exhibit overlapping phenomena between two or more classes, resulting in certain classification errors. In comparison to the aforementioned classification methods, the proposed PFA-CNN method in this study demonstrates excellent performance in recognizing test samples. It achieves high classification accuracy for each category of clothing with minimal overlapping, and only a very small number of samples are incorrectly classified. Therefore, the FPA-CNN method proposed in this paper effectively classifies clothing from different regions of the She ethnic group and achieves higher classification accuracy than the SVM, BP, RNN, and RBF. This verifies that the proposed method has higher accuracy and efficiency.

3.3.4. Comparative Analysis of Classification Methods

To validate the superiority of the proposed method in this paper, we compared it with the CNN model using adaptive feature extraction. In order to ensure consistency in the classification model parameters, the experiments were conducted with the same parameters as described in Section 3.2.1, and the comparative experimental classification results are presented in Table 4.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Training Time/s</th>
<th>Average Classification Accuracy/%</th>
<th>Highest Classification Accuracy/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>25.14</td>
<td>91.94</td>
<td>93.24</td>
</tr>
<tr>
<td>Proposed method</td>
<td>12.15</td>
<td>98.38</td>
<td>98.65</td>
</tr>
</tbody>
</table>

According to Table 4, the training time for classifying She ethnic clothing using the CNN classification model is 2.07 times longer than that of the proposed method in this paper. Moreover, both the average and highest classification accuracies are lower compared to the results of the proposed method. This indicates that the utilization of the FPA and kernel principal component analysis for optimizing feature parameters effectively reduces the impact of redundant information on the recognition accuracy of the CNN, thereby significantly improving its recognition performance.

4. Discussion

The experimental section of this study primarily focuses on She ethnic clothing color extraction, clothing classification, and a comparison of classification methods. Firstly, in the color extraction phase, we unearthed the fuzzy relationship between the distribution of She ethnic groups in different regions and the combination of clothing colors. We discussed the characteristics of clothing among different branches of the She ethnic group. Secondly, this study achieved an overall classification accuracy of 98.65% by constructing a neural network discriminative model tailored to the distinctive features of She ethnic clothing. For misclassifications, these were mainly attributed to the small proportion of color features in certain patterns, preventing the formation of effective feature distinctions. Additionally, feature omissions could be present in the process of image collection and reconstruction. Finally, in the comparative experiments of clothing classification methods, the constructed model exhibited significant effectiveness in She ethnic clothing classification. This is because, compared to traditional classification algorithms, this experiment employed receptive fields and weight-sharing techniques, reducing the number of required network weights and burdens during training. Furthermore, through supervised learning, the model’s information acquisition capability was further enhanced, effectively improving classification accuracy.

The study still has certain limitations, such as a scarcity of datasets. Datasets are essential for training models, yet the existing collection of She ethnic clothing images falls short of the required quantity, which necessitates sourcing traditional clothing images from the internet or museums. The collected images also require a series of preprocessing...
steps before they can be used. As of now, there is no well-established standardized dataset available.

In future research, the focus will continue to be on the recognition and classification of ethnic clothing. Firstly, efforts can be directed towards improving classification accuracy. Although the FPA-CNN method proposed in this paper yielded favorable results in classifying She ethnic clothing, there is still room for improvement. Further optimization of feature extraction and classification models, such as introducing additional feature parameters, can enhance classification accuracy. Secondly, there will be an expansion of the variety of ethnic clothing types. In China, there are other precious ethnic clothing cultures that are on the verge of being forgotten, such as Tibetan and Dai ethnic clothing. The restoration and recognition of clothing images from these ethnic groups hold significant practical implications for the preservation of intangible cultural heritage.

5. Conclusions

This study focuses on the She ethnic clothing from five regions in China: Jingning, Fu’an, Luoyuan, Xiapu, and Fuding. This study extracts color histograms and color moment features and applies FPA optimization to fuse the color features. Then, using kernel principal component analysis and a CNN deep learning framework, the study classifies the She ethnic clothing from the five regions. The main conclusions are as follows:

(1) A collection of 265 color histograms and color moment feature vectors are extracted from the clothing images of the five branches of the She ethnic group, revealing the color feature distribution of each branch. Black and blue are the dominant colors in She ethnic clothing, accounting for a range of [55.91%, 96.65%], with the Jingning branch having the highest proportion. The complementary colors mainly include shades of pink and orange-red, accounting for a range of [3.35%, 44.09%], with the Luoyuan branch having the most diverse color palette.

(2) Through the FPA, the optimal weight fusion is obtained. It is found that the highest classification accuracy for the FPA color feature fusion is achieved when the weights for the color histograms and color moment features are 0.645 and 0.355, respectively, resulting in an accuracy of 97.29% with a training time of only 13.25 s.

(3) The proposed FPA-CNN model, which incorporates receptive fields and weight-sharing techniques, achieves an average classification accuracy of 98.38% for She ethnic clothing. Compared to the SVM, BP, RNN, and RBF models, the FPA-CNN model improves the accuracy by 11.49%, 7.7%, 6.49%, and 3.92%, respectively.

Author Contributions: Conceptualization, X.D. and F.Z.; methodology, X.D. and T.L.; software, X.D.; validation, X.D. and T.L.; formal analysis, X.D. and T.L.; investigation, J.C.; resources, X.D. and J.C.; writing—original draft preparation, X.D.; writing—review and editing, X.D. and T.L.; visualization, X.D. and L.M.; supervision, F.Z.; project administration, X.D. and F.Z.; funding acquisition, X.D., T.L., J.C., L.M. and F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Laboratory of Ministry of Culture and Tourism Foundation of China (NO. 19076223-B), the Clothing Engineering Research Center of Zhejiang Province (NO. 2019FZKF08), and the Zhejiang Provincial Philosophy and Social Sciences Planning Project (NO. 22NDJC077YB).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available on request from the authors.

Acknowledgments: The authors would like to thank the anonymous reviewers for their insightful comments and constructive suggestions.

Conflicts of Interest: The authors declare no conflict of interest.
References


33. Khan, U.A.; Javed, A. A hybrid CBIR system using novel local tetra angle patterns and color moment features. *J. King Saud. Univ.-Com.* 2022, 34, 7856–7873. [CrossRef]
50. Kaya, E. Quick flower algorithm (QFPA) and its performance on neural network training. *Soft Comput.* 2022, 26, 9729–9750. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.