Journal of Network Intelligence ©2024 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **9**, Number **2**, May **2024**

Research on Oil Painting Style Classification Based on Spatial Reduction and Fuzzy Clustering

Xiao-Hong Heng*

Chuzhou City Vocational College,
Chuzhou 239057, P. R. China
15955069405@163.com

Zong-Fei Liu

School of Information Technology and Engineering,
St. Paul University Philippines,
Tuguegarao 3500, Philippines
xa3506@163.com

*Corresponding author: Xiao-Hong Heng
Received October 18, 2023, revised December 27, 2023, accepted March 15, 2024.

Abstract. *The existing oil painting classification algorithms can not efficiently classify the style of oil painting. Aiming at this problem, this paper introduces multi-label and genetic factors, and designs an oil painting style classification algorithm based on spatial dimension reduction and fuzzy clustering. Firstly, the algorithm conducts principal component analysis on the oil painting data set, separates the correlation of low-dimensional features, and introduces the label factor through PCA to project high-dimensional features into the low-dimensional space to retain a large amount of effective information while reducing dimensionality, thus achieving the purpose of dimensionality reduction and noise reduction in the feature space. Then, adaptive genetic algorithm space is used to cluster the dimensionality reduction of oil paintings. The calculation method of genetic variation probability is improved, and the multi-label and genetic factors are searched and classified globally by using cluster iteration formula. Finally, in order to verify the performance of the algorithm, this paper conducts simulation experiments based on MATLAB. The experimental results show that when the genetic factor is 0.8, the accuracy rate, precision rate and recall rate of the proposed algorithm are 92.48%, 82.63% and 95.61%, respectively. Compared with other algorithms, the accuracy rate, precision rate and recall rate of oil painting classification are effectively improved.*
Keywords: spatial dimension reduction; fuzzy clustering; multi-label classification; global search; genetic variation

---

1. **Introduction.** With the rapid development of artificial intelligence, people can use various search tools to retrieve all kinds of image information on the network. How to classify all kinds of images accurately has become the research direction of scholars. Traditional image classification studies are mainly fine-grained classification based on image content, focusing on color, shape, object layout, etc. [1, 2, 3]. Common oil painting styles include expressionism, Fauvism, post-impressionism and Cubism [4]. The classification of oil painting art style can not only provide supplementary information for art appreciation, but also provide ideas for future oil painting creation. However, the existing image classification technology is not practical for the classification of oil painting style, and the classification of oil painting is mostly completed through the subjective judgment

of scholars, which is easy to ignore the details and lead to wrong classification [5, 6, 7]. Subsequently, more scholars used deep learning algorithms to automatically extract and classify artworks. Because the definition of artistic style is subjective and lacks specific quantitative indicators, how to classify oil painting style accurately is a very challenging work in terms of oil painting style classification.

1.1. **Related Work.** The research on oil painting style classification algorithm is mainly divided into two kinds. One is based on manual feature extraction. The other is based on deep learning. To solve the single problem of extracting the style features of painting images, Zhou et al. [8] and Zeng et al. [9] proposed a series of effective algorithms. However, they have strict restrictions on the input images, which must be gray scale images, so the styles of works are required to be very different. In view of the problem that input images needs to be grayscale images, Peng et al. [10] used image palette entropy to analyze the artist's oil painting creation and study the characteristic value of paintings, thus discovering the difference and connection between the creation of different artists. An algorithm was proposed by Yang et al. [11] on the ground of information entropy to calculate block entropy, color entropy and contour entropy of images according to the relationship between colors, line contours, and the area ratio of painting targets and backgrounds among works of different styles, which is used for the style classification of paintings. Aiming at the problem that image quality has a great influence on style classification effect. Hu et al. [12] using Gabor energy, conducted comparative experiments and proved that the higher the image quality, the more conducive it is to style classification. However, low-resolution images still have the ability to represent style. When the number of images is insufficient, low-quality images can be selected for experimental research. Wang and Huang [13] proposed an oil painting art style description method based on key regions, aiming at the problem that the features extracted by the existing oil painting feature extraction algorithm could not accurately describe the artist's art style, and applied it to oil painting classification.

With the rapid popularization of computer tools, easy access to image data, sufficient number of images and easy access to computer hardware, most of the current research tends to adopt deep learning algorithms. Aiming at the automatic recognition of digital painting art, Krinidis and Chatzis [14] proposed an oil painting classification system based on color features to isolate misclassified images and gradually correct errors. Lamnerti et al. [15] proposed an unsupervised technique for automatically extracting individual brushstroke characteristics and applied it to the identification and classification of oil paintings. A cropped CNN model was presented by Du et al. [16]. It extracted style features and classify oil painting styles. On the ground of IBS, ModIbs distance was defined to evaluate the similarity between different artist works. Sun et al. [17] trimmed the weight of the model for the purpose of shorting the training time of the model, and the accuracy reached 51.5%. Li et al. [18] proposed a brushstroke feature extraction method on the ground of edge detection and cluster segmentation, which was used to distinguish Van Gogh's paintings in different development periods.

Most of the above methods use the type of single feature, only describe the details of oil painting, can not describe the overall style of oil painting, so it is not suitable for the extraction of oil painting style. On the ground of human thinking habits, Zhang et al. [19] established descriptors in accordance with the similarity principle and feature vector, for the purpose of representing the fashion features of visual art images, and classified visual art works' style by extracting artistic style features. Different painting style parts are designed. Secondly, according to the different painting skills used by the painter, the different understanding of light and shadow lead to the different brightness of the portrait, but there is a problem of low classification accuracy.

1.2. **Motivation and contribution.** For the purpose of improving the accuracy and efficiency of the above classification algorithms, this paper proposes an oil painting style classification algorithm based on label space reduction and fuzzy clustering. This algorithm first conducts principal component analysis on oil painting data set to separate the correlation of low-dimensional features, so as to achieve the purpose of dimensionality reduction and noise reduction in feature space. Then, a multi-label classification method based on feature label reduction is designed. On this basis, adaptive genetic algorithm space is introduced to cluster oil paintings with dimensionality reduction and improve the calculation method of variation probability. The fitness of genetic algebra and classification labels is improved by using cluster iteration formula. The experimental results show that the proposed algorithm can solve the clustering center more accurately, and effectively improve the accuracy, accuracy and recall rate of oil painting classification.

2. Basic theoretical knowledge.

2.1. **Fuzzy C-Means clustering algorithm.** The Fuzzy C-Means clustering algorithm (FCM) can automatically classify samples [20, 21, 22, 23]. Scholar Bezedek demonstrated the convergence of the FCM algorithm and established a corresponding theoretical system based on the research progress of the FCM algorithm [24]. Subsequently, FCM theory was applied to the field of image processing. Unlike the K-means algorithm, which is a representative algorithm of hard clustering algorithms, the FCM clustering algorithm allows for membership values between 0 and 1, making pixel classification more flexible and reliable.

The basic idea of the FCM algorithm is to use the Lagrange multiplier method to obtain the minimum value of the objective function, then iteratively solve to obtain the optimal clustering center and membership, and finally divide the pixels. The expression of its objective function is shown in Equation (1):

$$\tau(\mu,\nu) = \sum_{i=1}^{m}\sum_{j=1}^{n}\mu_{ij}^{t}c^{2}(\chi_i,\nu_j) \tag{1}$$

Satisfy the Equation (2):

$$0 < \mu_{ij} < 1, \sum_{j=1}^{n}\mu_{ij} = 1 \tag{2}$$

Among them, m is the number of pixels, n is the number of cluster centers, μ is the membership degree matrix $m * n$ of size, ν is a set containing n cluster centers, μ_{ij} is the membership degree of pixels χ_i to cluster centers ν_j, $c^2(\chi_i,\nu_j)$ is the Euclidean distance between pixels χ_i and cluster centers ν_j, t is the fuzzy index, generally taken as 2.

Minimize the objective function of the FCM algorithm through Lagrange number multiplication, and define a Lagrange function by linking the objective function with the following formula:

$$L = \sum_{j=1}^{n}\sum_{i=1}^{m}\mu_{ij}^{t}c^{2}(\chi_i,\nu_j) + \sum_{i=1}^{m}\delta_i(\sum_{j=1}^{n}\mu_{ij} - 1) \tag{3}$$

Take the partial derivatives of the above equations for μ_{ij} and ν_j, and set them to 0, to obtain the following formula:

$$\frac{\partial L}{\partial \mu_{ij}} = 0 \tag{4}$$

$$\frac{\partial L}{\partial \nu_j} = 0 \tag{5}$$

Thus, we can obtain Equation (6) and Equation (7):

$$t\mu_{ij}^{t-1}c^2(\chi_i, \nu_j) + \delta_i = 0 \tag{6}$$

$$2\sum_{i=1}^{m} \mu_{ij}^{t}(\nu_j - \chi_i) = 0 \tag{7}$$

By simplifying Equation (6), we can get μ_{ij}:

$$\mu_{ij} = (\frac{-\delta_i}{tc^2(\chi_i, \nu_j)})^{\frac{1}{t-1}} \tag{8}$$

Due to the existence of constraints $\sum_{\ell=1}^{n} \mu_{\ell j} = 1$, substituting the above Equation (7) into the left equation can be obtained δ_i:

$$\delta_i = -\frac{1}{\sum\limits_{k=1}^{n} \frac{1}{tc^2(\chi_j, \nu_k)}} \tag{9}$$

The Equation (8) and Equation (9) are the iterative formulas of membership matrix and cluster center respectively. When using Lagrange number multiplication to minimize the objective function, the membership matrix and cluster center need to be iterated continuously according to the Equation (6) and Equation (7) until the objective function is minimized. When the value of the objective function meets the convergence condition or the algorithm reaches the maximum number of iterations, the algorithm ends.

2.2. **Multi-label classification algorithm.** The fundamental idea of this algorithm is to transform the multi-label learning issue into the label sorting issue, and to use the pair comparison technology to sort the labels [25].

For w possible class labels $x_1, x_2, \ldots, x_w$, a total of binary classifiers (x_i, x_j) $(1 \leq i < j \leq w)$ can be obtained by pairwise comparison $w(w-1)/2$ of each label pair.

For each pair of labels (w_i, w_j), first, by considering the relative correlation of each training sample to w_i and w_j, the sample is then broken down into p separate binary classification subproblem, where each binary classification subproblem corresponds to all possible labels in the label space. For Class i labels χ_i, the algorithm considers the association between each sample and the label, thus establishing the binary training set as follows:

$$\begin{gathered} E_i = \{(y_j, \varphi(X_j, x_i)\ 1 \leq j \leq m\} \\ \varphi(X_j, x_i) = \begin{cases} +1, & if\ x_i \in X_j \\ -1, & otherwise \end{cases} \end{gathered} \tag{10}$$

A binary learning algorithm α is used to introduce a binary classifier f_j. For each multi-label training sample (v_j, X_j), instance y_j will participate in p binary classifier learning steps. For the associated label $x_i \in X_j$, x_i is a positive instance sample through $f_j(\cdot)$; On the other hand, for the unassociated label $x_l \in \overline{X_j}$, y_j is treated as a negative instance sample.

For sample y of unknown labels, the algorithm forecasts its related label set X by querying the label correlation on each distinctive binary classifier, and then merges all the associated labels:

$$X = \{x_i | f_i(y) > 0, 1 \leq i \leq p\} \tag{11}$$

3. **Multi-label classification algorithm based on spatial dimension reduction.** In multi-label classification, problems such as high feature dimension and sparse label space will be faced, and the direct use of traditional multi-label classification algorithms will affect the classification effect. Because of the collinear relationship between oil painting feature attributes, the generalization ability of the model will be weak, and the sparsity of the label space will make it difficult for the model to find data features. Aiming at the above problems, this paper proposes a space Encoding-Based Multi-Label Classification algorithm (SEMC algorithm) based on dimensionality reduction. Firstly, high-dimensional features are projected into low-dimensional Spaces by PCA method. A large amount of effective information is retained while dimensionality is reduced to prepare for the next step of fuzzy clustering, especially for large-scale multi-label data sets. Figure 1 shows the algorithm framework for this chapter.

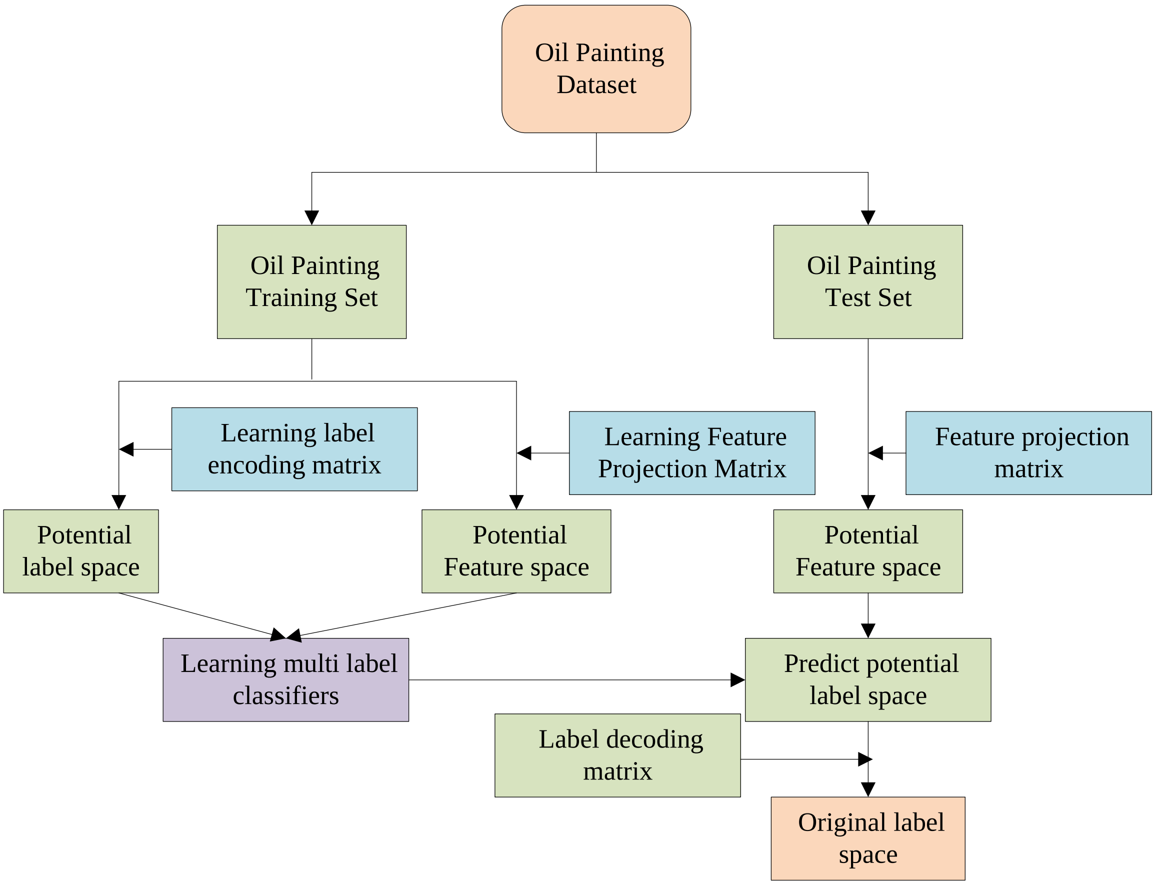

Figure 1. The frame of the Multi-label classification algorithm

3.1. **Feature space dimensionality reduction.** In the feature dimensionality reduction, in order to avoid overfitting, the training data set, the unsupervised feature dimensionality reduction PCA algorithm is used here. PCA is a classical feature extraction method, which can find a projection direction containing a large amount of effective information of data, and its calculation Equationis as follows:

$$\theta^2 = \frac{1}{m}\sum_{j=1}^{m}(\beta^T\gamma_j)^2 = \beta^T(\frac{1}{m}\sum_{j=1}^{m}\gamma_j\gamma_j^T)\beta = \beta^T\zeta\beta \tag{12}$$

In general, the eigenvector corresponding to the maximum eigenvalue of eigencovariance β is taken to form the front projection matrix β^* of the optimal moment ζ. Then, the low dimensional effective feature space, $Y' = Y\beta^*$ is obtained by using the projection matrix β^*.

3.2. **Label space coding.** Suppose the label matrix Ψ is already centralized. The label space coding method designed in this chapter consists of two parts, one is coding loss and the other is dependency loss. In the coding loss, the effective projection matrix is designed as follows:

$$F_f(\theta(\sigma(x_i))_{i=1}^m, X) = \| \tilde{X} - X \|_2^2, \nu^\tau \nu = I \tag{13}$$

where $\nu \in \mathbb{R}^{p\times t}$ is a linear projection matrix, t is the expected dimension of the low-dimensional label space. Matrix $\varpi = \delta(X) = X\nu$ represents the coding matrix.

Matrix $\tilde{\varpi}$ reconstruction matrix that represents an encoding matrix is $\tilde{\varpi} = \delta(X) = X\nu^T$. In the dependency loss section, this chapter trains a regression model, choosing directly from the feature space Y to the low-dimensional label space X, and then learns a classifier, so the second part of multi-label space coding is the Equation (14).

$$F_f(X, Y') = (m-1)^{-2} tr(\alpha\beta Q\beta) \tag{14}$$

where β is a kernel matrix.

By combining the Equation (13) and Equation (14), the resulting global optimization objective function is as follows:

$$\min_{\nu} F(\nu) = \|X\nu\nu^T - X\|_2^2 - atr(\alpha\beta Q\beta), \quad \nu^T\nu = 1 \quad (15) \tag{15}$$

where α is the regularization parameter of balance dependent loss and coding loss, this paper adopts linear kernel for Q, that is, $Q = XX^T = Z\nu\nu^T Z^T$. The kernel matrix α of the sample is the Equation (16).

$$\alpha_{ij} = \begin{cases} \exp\left(\dfrac{-\|y_i' - y_j'\|_2^2}{2\vartheta^2}\right), \ if\ y_i' \in M_z(y_j')\ or\ y_j' \in M_z(y_i') \\ 0,\ otherwise \end{cases} \tag{16}$$

Where $\theta = 1$, $Z = 10$, $M_z(y_j')$ indicates the y_i' nearest neighbors of Z. Z is a symmetric matrix that takes advantage of local information and needs to search z nearest neighbors for each sample.

The Equation (15) can be converted to the following:

$$\begin{aligned} &\min_{\nu\nu^T=1} F(\nu) = \|X\nu\nu^T - X\|_2^2 - atr(\alpha\beta Q\beta), \\ &\Leftrightarrow \min_{\nu\nu^T=1} \|X\nu\nu^T - X\|_2^2 - atr(\alpha\beta Xvv^T X^T\beta), \\ &\Leftrightarrow \min_{\nu\nu^T=1} tr(v^Tvv^TX^TvvQ - 2v^TX^TXv) - atr(\alpha\beta Xvv^TX^T\beta) \\ &\Leftrightarrow \max_{\nu\nu^T=1} tr(v^TX^T(1 + a\beta\alpha\beta)Xv), \end{aligned} \tag{17}$$

The final problem is converted to solving the eigenvalue of $B = X^T(1 + a\beta Q\beta)X$. Since B is a heaped matrix, its eigenvalues are all real numbers. Let the eigenvector $v_1, \ldots, v_m$ correspond to the eigenvalue $v_1, ..., v_m$ of the matrix B, then the corresponding optimal label projection matrix is $v^* = [v_1, \ldots, v_n]$. The optimization problem is solved by solving the eigenvector of $m \times m$ matrix.

Therefore, the algorithm designed in this paper encodes feature space and label space step by step, which is suitable for multi-label classification problems with a large number of class labels and features. First, two sets of low-dimensional effective spatial representations are found, and then the adaptive model is learned on the processed low-dimensional space to label various styles of oil painting, which lays a foundation for the research of canvas style classification method.

4. **Oil painting style classification method on the basis of spatial reduction and improved adaptive clustering algorithm.**

4.1. **Algorithm running parameters.** Based on the research of multi-label classification algorithm with spatial dimensionality reduction described above, considering the influence of prior information of oil painting images and the robustness and applicability of fuzzy clustering algorithm, this paper first introduces multi-label classification factors to determine the initial clustering center of the clustering algorithm, and then performs global search through genetic algorithm to improve the accuracy and recall of classification. The running parameters of the clustering algorithm will greatly affect the convergence and efficiency of the algorithm, and the operating parameters of the algorithm are as follows:

Classification factors: $\alpha = 0, \beta = 1$.

Population size M: Population size refers to the number of individuals in the population, generally between $30 \sim 75$ is more appropriate.

Evolutionary algebra F: Evolutionary algebra is the maximum number of iterations of the algorithm preset in advance, and the general value is $100 \sim 500$.

Crossover probability q_c: The crossover probability is too small, and the search range becomes larger, resulting in low algorithm efficiency; If the crossover probability is too large, it will lead to the destruction of the better performing individuals. Generally, it is more appropriate to take between $0.4 \sim 0.9$.

Mutation probability q_m: The probability of mutation is too small, which will affect the diversity of the population; If the probability of mutation is too large, it will lead to the destruction of the better performing individuals. The probability of variation is generally taken as $0.01 \sim 0.1$.

4.2. **Oil painting style classification method based on spatial dimensionality reduction and improved self-adaptive clustering algorithm.** Therefore, this paper introduces genetic algorithm and multi-label classification factors to determine the initial cluster center, first of all, aiming at the problem that the traditional genetic algorithm fixes the crossover, mutation probability and is prone to precocious phenomenon, and introduces the multi-label classification method as a selection operator. At the same time, according to the size and change speed of individual fitness in the population and the algorithm evolution process, the crossover probability and mutation probability are adaptively adjusted to avoid the algorithm falling into local extremes as much as possible. The specific algorithm flow of this chapter is shown in the Figure 2.

In this paper, genetic algorithm and multi-label classification factor are introduced as selection operators, and the following probability distributions can be obtained by descending and sorting the fitness of population individuals and corresponding numbers:

$$q(j) = p(1-p)^{j-1} \; j = 1, 2, ..., m \tag{18}$$

where p is variable parameters, j is the ordinal number of the sort. For a given different p, the following probability distribution function can be obtained, as shown in Figure 3.

By relatively increasing the value of the probability of crossing and variation, the probability of their elimination in the process of evolution is increased. The probability of destruction of highly suitable individuals and the probability of retention of low-fitness

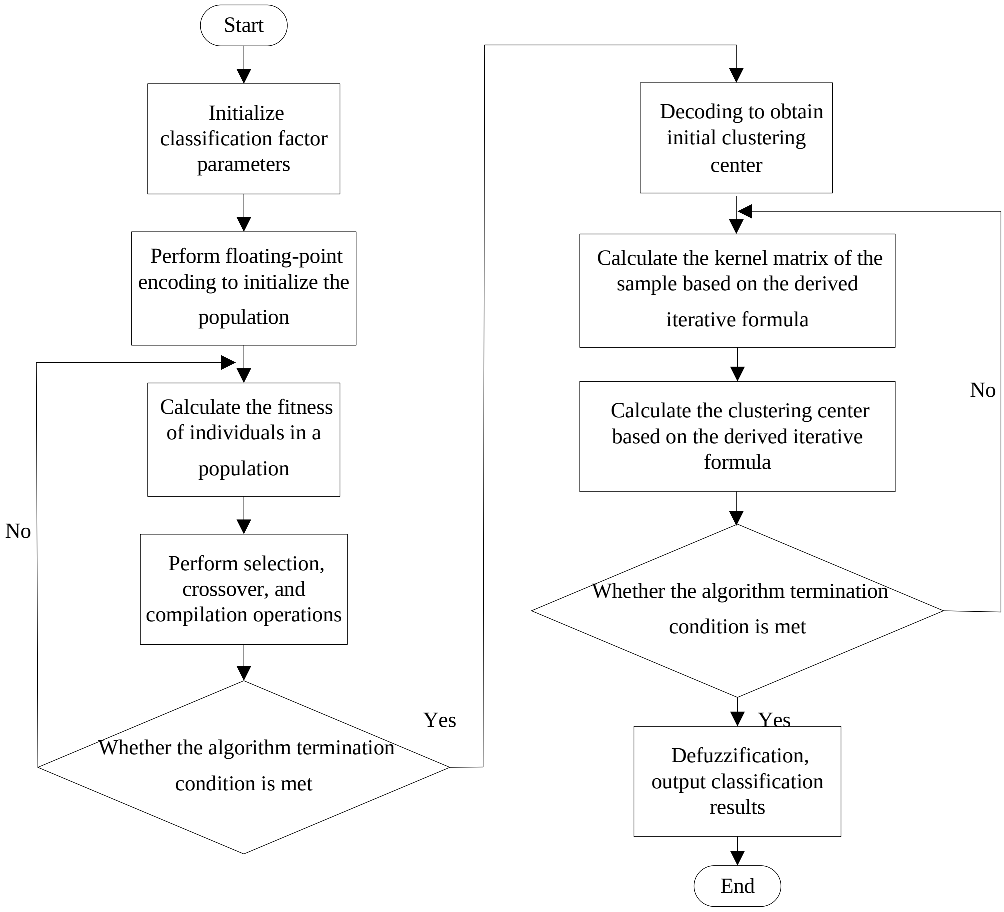

Figure 2. The flowchart of the algorithm

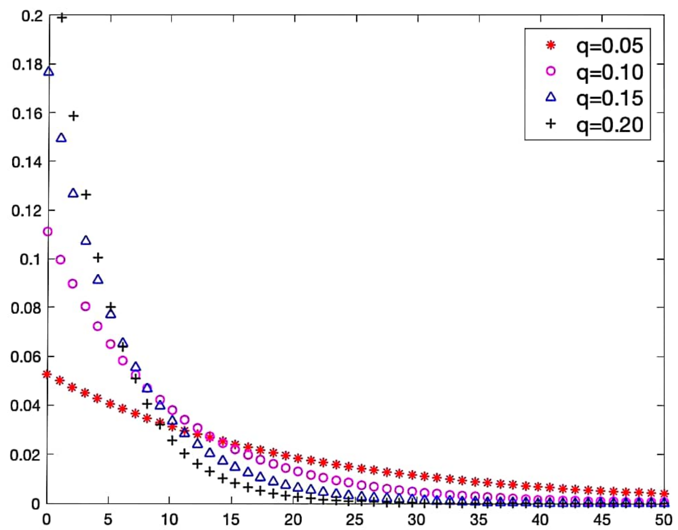

Figure 3. Probability distribution curve

individuals can be suppressed, and the improved calculation Equation for crossover and mutation probability is as follows.

$$q_c(s,i) = \begin{cases} q_c(s)\dfrac{F_{\max} - F_i}{F_{\max} - F_{avg}} \ F_{avg} \le F_i \\ q_c(s) \ F_{avg} \le F_i \end{cases} \tag{19}$$

$$q_n(t,i) = \begin{cases} q_n(s)\dfrac{F_{\max} - F_j}{F_{\max} - F_{avg}} \ F_{avg} \le F_j \\ q_n(s) \ F_{avg} \le F_j \end{cases} \tag{20}$$

In the above formula, $F_{\max}$ is the maximum fitness of individuals in the s generation population, F_{avg} is the average fitness of individuals, F_i is the fitness value of the i pair of individuals with crossover, and F_j is the fitness value of individual j with variation.

For the purpose of increasing crossover probability and reducing mutation probability, a variation operation was used to increase population diversity. Therefore, the sum is shown in the Equation (20) and (21).

$$q_c(s) = q_c - a\frac{s}{S} \quad (21) \tag{21}$$

$$q_n(s) = q_n + a\frac{s}{S} \quad (22) \tag{22}$$

Where q_c represents the initial crossover probability, q_n represents the initial mutation probability, s represents the algebra of current evolution, S represents the total evolutionary algebra, and a is the label classification factor in the previous section. After improvement, the crossover probability $q_c(s)$ will gradually decrease, while the mutation probability $q_n(s)$ will gradually increase.

Then, for the objective function of $q_c(s)$, the extreme value is obtained, and the Lagrange function is constructed under the constraint condition $1 - \sum_{i=1}^{c} v_{ic} = 0$, as shown in Equation (23).

$$L = \sum_{i=1}^{m}\sum_{c=1}^{n}[v_{ci}^{m}||y_j - v_c||^2 + F_{ci}] + \sum_{j=1}^{n}\lambda_j(1 - \sum_{i=1}^{c} u_{ic}) \tag{23}$$

Plug in the classification factor expression and take the partial derivative of v_{ci} and u_c respectively and set them to 0.

$$\frac{\partial L}{\partial v_{ci}} = nv_{ci}^{n-1}||y_j - u_c||^2 - n(1 - v_{ci})^{n-1}||y_j - u_c||^2 \sum_{i\in M, i\neq j}\frac{1}{1 + c_{ij}} - \lambda_j = 0 \tag{24}$$

Under constraint condition $\sum\limits_{i=1}^{c} v_{ic} = 1$, the Equation (24) can be obtained.

$$\sum_{c=1}^{j}\frac{\frac{\lambda_j}{2||y_i - u_c||^2} + \sum\limits_{i\in M, i\neq j}\frac{1}{1+f_{ij}}}{1 + \sum\limits_{i\in M, i\neq j}\frac{1}{1+f_{ij}}} = 1 \tag{25}$$

Based on the above formula, the Lagrange multiplier λ_i can be obtained.

$$\lambda_i = 2\frac{1 + (1 - c)\sum\limits_{i\in M, i\neq j}\frac{1}{1+f_{ij}}}{\sum\limits_{i=1}^{c}\frac{1}{||y_j - u_c||^2}} \tag{26}$$

Finally, the classification equation of cluster center can be obtained.

$$u_c = \frac{\sum_{j=1}^{m} \left(v_{ic}^n y_i + \sum_{i \in M, i \neq j} \frac{(1-v_{ic})^n}{1+f_{ij}} y_i\right)}{\sum_{j=1}^{m} \left(v_{ic}^n + \sum_{i \in M, i \neq j} \frac{(1-v_{ic})^n}{1+f_{ij}}\right)} \tag{27}$$

5. Simulation and evaluation.

5.1. Parameter Settings.

In order to verify the effectiveness of the algorithm proposed in this paper, data sets with high classification accuracy in existing studies are selected as the data sets of this paper, as shown in Table 1.

Table 1. Data Set.

	Classicalism	Surrealism	Training set	Test data
Data set1	67	162	420	139
Data set2	651	360	190	157
Data set3	263	284	1260	261
Data set4	459	635	486	208

For the classifier used in this paper, a multi-label classification algorithm on the ground of spatial coding dimensionality reduction is used to calculate its parameters, and the search interval is set to $[0, 40]$. The size of the four ordinary oil painting images is $512*512$, while the image size in the oil painting dataset is $100 * 100$. In the experiment, set the population size of the clustering algorithm $M = 30$, the initial crossover probability $p_s = 0.6$, the initial mutation probability $P_t = 0.1$, the maximum number of iterations $w = 200$, $\theta = 10^{-7}$ and the corresponding $a = 0.4$, $\beta = 0.2$, $\mu = 1.4$. Set $m = 2$, $n = 4$, $v = 200$, $a = 1$ in the label classification algorithm. For the purpose of asserting the robustness of the algorithm, the training set was randomly classified by a ratio of 6:4 in the experiment of common oil painting images, and each set of parameters was evaluated by cross-validation. The parameters with the highest average accuracy were run 100 times under the same data set, and the parameters with the highest average accuracy were taken as the final classifier parameters, as shown in Table 2.

Table 2. Values of classifier parameter

Classifier	Parameter	value
SVM	Coefo	0.6
RF	poly	0.8
CNN	Sigmoid	1

5.2. Experimental design and result analysis.

In this paper, four common oil painting classification algorithms are tested on oil painting data set. For the convenience of description, the algorithm designed in this paper is denoted as Ours, the algorithm in literature [26] is denoted as Ism, the algorithm in literature [27] is denoted as Acs, the algorithm in literature [28] is denoted as Aig, and the algorithm in literature [29] is denoted as Lta. From Table 3, it can be seen that Ism and Aig is of poor style classification effect on oil painting data set, in which the macro average of Aig (weighted average of accuracy, accuracy and recall rate) is 70.3%, while the macro average of Ism is slightly better, with 75.56%. The average macro of Acs on the oil painting data set is 76.01%.

Compared with other oil painting classification algorithms, the algorithm of Lta series has a relatively better recognition effect on the oil painting data set. The classification effect of Lta on oil painting dataset is 82.63%. Next, this paper improves the model recognition effect of the improved classification algorithm.

Table 3. Comparative Experiments of Different Algorithms on Oil Painting Datasets

Number	classification method	Macro AveLtae
1	Ism	75.56%
2	Acs	76.01%
3	Aig	70.3%
4	Lta	82.63%

Multi-label classification factor is introduced into the above four classification algorithms, and the key features of oil painting images are captured first, and the genetic mechanism is embedded on this basis, and then the recognition effect is compared by experiments respectively. It can be seen from Table 4 that the effects of the four models are improved after the introduction of classification algorithms or genetic mechanisms.

Table 4. Comparative Experiments of Enhanced Different Algorithms on Oil Painting Datasets

Number	classification method	Macro AveLtae
1	Ism+inheritance	78.83%
2	Acs+label	80.51%
3	Aig+label	73.05%
4	Lta+inheritance	87.46%

Therefore, the confusion matrix of the predicted results on the oil painting data set after the improved algorithm in this paper is shown in Figure 4. In the confusion matrix, the horizontal coordinate is the prediction label and the vertical coordinate is the true label. Of the Cubist images, 73% were correctly judged to be Cubist, 5% were incorrectly predicted to be realistic, 4% were incorrectly predicted to be brutalist, 1% were incorrectly predicted to be impressionist, and 2% each were predicted to be Baroque, romantic, and Rococo. From the confusion matrix, it can be seen that the improved algorithm can predict more than 50% of all the styles of images in the oil painting data set.

Next, we compare the improved classification algorithm with other four types of algorithms proposed by predecessors, namely Ism, Acs, Aig and Lta, on the oil painting data set. The classification outcome is listed in Table 5. It lists the comparison between the Accuracy (Acc), Precision (Prec) and Recall (Rec) of the four algorithms using SVM classifier and the algorithm designed in this paper (Ours). As can be seen from TABLE 5, Acc of Acs is 81.44%, which is 2.71% higher than Ism, 5.6% higher than Prec, and 6.96% lower than Rec. Acc, Prec, and Rec of Lta are 10.87%, 16.54%, and 9.62% higher than Aig. Compared with Ism, AcS, Aig and Lta, Ours algorithm designed in this paper has significantly improved in Acc, Prec and Rec, and can classify oil painting styles more accurately.

In order to more strictly verify the robustness of Ours algorithm, this paper immediately divides the oil painting data set into 10 pieces without repeating them, and selects one of them for marking as the test set. After the training, the test set is used for testing, and

	abstract expressionism	cubism	fauvism	impressionism	romantism	realism	Braoque style	Rococo
abstract expressionism	0.53	0.06	0.0	0.03	0.08	0.01	0.0	0.0
cubism	0.05	0.61	0.06	0.0	0.0	0.16	0.05	0.09
fauvism	0.12	0.02	0.54	0.06	0.08	0.29	0.12	0.15
impressionism	0.0	0.01	0.05	0.65	0.31	0.01	0.24	0.17
romantism	0.03	0.11	0.01	0.09	0.75	0.06	0.01	0.0
realism	0.06	0.14	0.14	0.11	0.15	0.93	0.0	0.01
Braoque style	0.18	0.04	0.02	0.17	0.01	0.0	0.58	0.01
Rococo	0.26	0.54	0.08	0.01	0.0	0.01	0.0	0.64

Figure 4. Confusion matrix

Table 5. Results of different classification method

Method	Acc	Prec	Rec
Ism	78.73%	58.39%	89.57%
Acs	81.44%	63.99%	82.61%
Aig	73.28%	53.67%	83.9%
Lta	84.15%	70.21%	93.52%
Ours	93.61%	87.36%	95.87%

the above steps are repeated for 10 times, so that each piece of data can be used as the test set for a test. The average and optimal values of the 10 experimental results were calculated, and the results were compared, as shown in Table 6. The overall evaluation effect of the fifth experiment is the best, Acc, Prec and Rec are 95.37%, 90.15% and 97.2% respectively, and the average value of the 10 experiments is 93.44%, 89.04% and 95.78% respectively, which is little different from the optimal result, indicating that Ours algorithm has good robustness.

Table 6. Evaluation values

	Acc	Prec	Rec
Experiment1	94.28%	88.51%	95.12%
Experiment2	92.0%	89.26%	93.05%
Experiment3	94.35%	88.1%	98.04%
Experiment4	91.09%	88.69%	95.27%
Experiment5	95.37%	90.15%	97.2%
Experiment6	92.74%	89.03%	98.47%
Experiment7	93.16%	88.27%	96.11%
Experiment8	92.04%	89.16%	95.1%
Experiment9	95.31%	89.08%	94.09%
Experiment10	94.06%	90.11%	94.38%
AveLtae value	93.44%	89.04%	95.78%

In this paper, label factor and genetic mechanism are added to Ours algorithm for classification experiment. The genetic factor ranges from 0.1 to 1, and is compared with Ism, Acs, Aig and Lta in three aspects of accuracy, accuracy and recall. As shown in

Figure 5-7, when the genetic factor is 0.8, Acc of Ism classification algorithm is 79.59%, Prec is 57.59%, and Rec is 86.3%, Acc of Acs classification algorithm is 84.52%, Prec is 62.52%, Rec is 82.52%, and Acc of Aig algorithm is 72.75%. Prec is 52.75%, Rec is 82.75%, Acc of Lta algorithm is 82.3%, Prec is 72.3% and Rec is 88.3%. Compared with Ism, Acs, Aig and Lta, Ours algorithm improves Acc by 12.89%, 7.96%, 19.73% and 10.18% respectively. Prec increased 25.04%, 20.11%, 29.78%, 10.23%, and Rec increased 8.02%, 13.09%, 12.86%, and 7.31%, respectively.

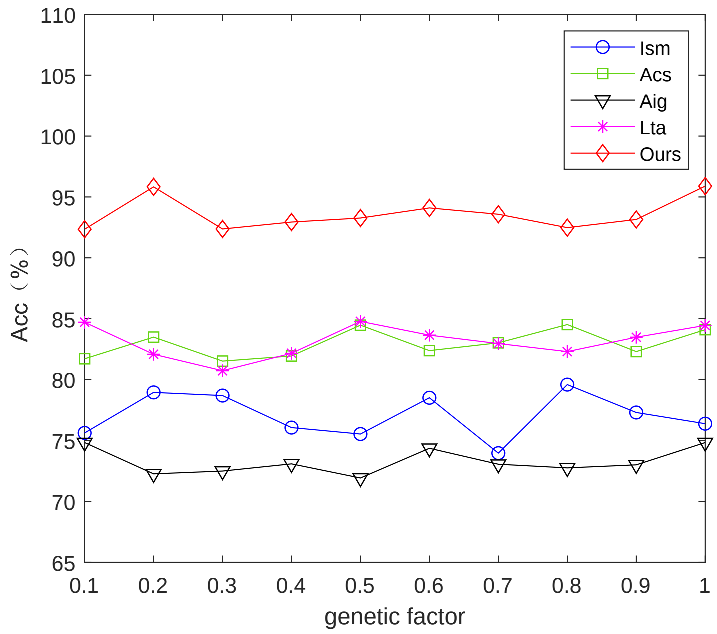

Figure 5. Accuracy rate

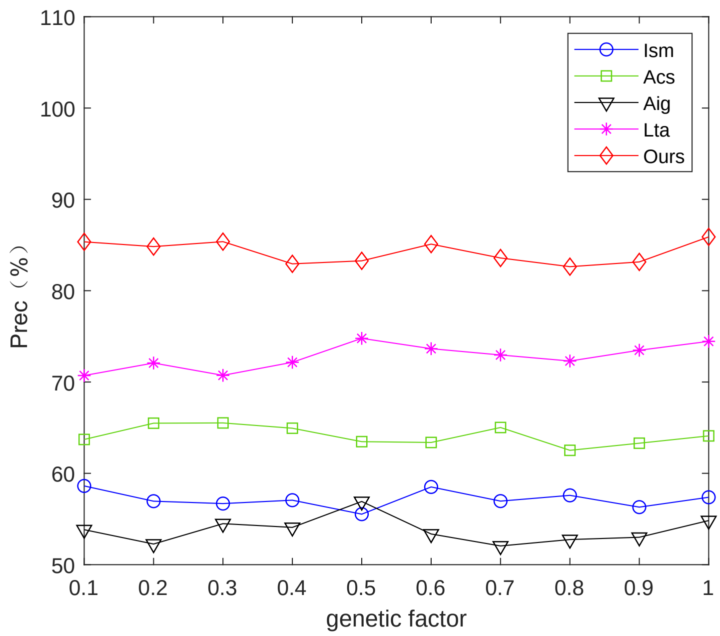

Figure 6. Precision rate

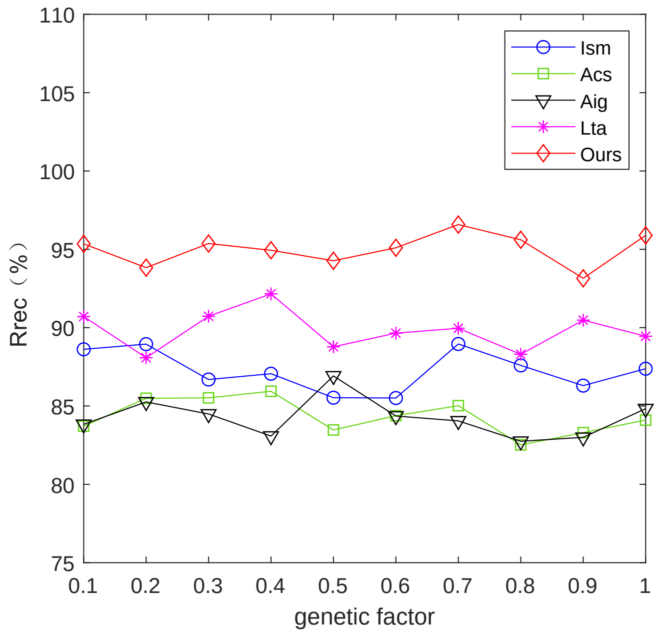

Figure 7. Recall rate

Since Aig and Ism only consider features such as the color of the oil painting and do not add an overall label to the classification to define its oil painting style, the classification effect is modest. However, Acs and Lta only use deep learning models to classify oil painting art styles, but the training set is too small and the classification effect is not good. The algorithm Ours designed in this paper is based on the principles of spatial reduction and fuzzy clustering. Starting from the label of oil painting itself, it uses genetic mechanism to measure the diversity, unity and balance of the picture, so as to distinguish the art styles of oil painting of different schools, thus achieving the best accuracy rate, precision rate and recall rate.

6. **Conclusion.** Based on the concept of space reduction and fuzzy clustering, this paper proposes an oil painting style classification algorithm based on label space reduction and fuzzy clustering. The algorithm first conducts principal component analysis on the oil painting data set, separates the correlation of low-dimensional features, and introduces label factors through PCA to project high-dimensional features into low-dimensional space, retaining a large amount of effective information while reducing dimensionality. On this basis, adaptive genetic algorithm space is used to cluster the oil paintings with dimensionality reduction, and improve the calculation method of variation probability. The global search classification of genetic algebra and label factors is carried out by using clustering iterative formula. The experimental results show that the proposed algorithm can solve the clustering center more accurately, and effectively improve the accuracy, precision and recall rate of oil painting classification. In the future work, this paper will integrate the painting features based on oil painting itself into the clustering algorithm to carry out further research.

Acknowledgment. This work is supported by Ministry of Education's first co-education project in 2019(No. 201901051025), Computer Basic Education Research Project of the National Institute of Higher Learning (No. 2019-AFCEC-044), Humanities and Social Science Project of Ministry of Education: Research on the Formation and Inheritance of the Characteristics of Eastern Anhui Folk Painting Project (No. 18YJA760021), Anhui Provincial Humanities and Social Sciences Project: Research and Development and Application of Cultural Tourism Derivatives of "Chuzhou Impression" Project (No. SK2021A0979).

REFERENCES

[1] C.-W. Deng, G. Wang, and J. Xu, "Piecewise two-dimensional normal cloud representation for time-series data mining," *Information Sciences*, vol. 374, pp. 32-50, 2016.

[2] C. Stewart, R. Montanaro, M. Sala, and P. Riccardi, "Feature Extraction in the North Sinai Desert Using Spaceborne Synthetic Aperture Radar: Potential Archaeological Applications," in *Remote Sensing*, vol. 8, no. 10, 825, 2016.

[3] W. Ahmad and D. Kim, "Estimation of flow in various sizes of streams using the Sentinel-1 Synthetic Aperture Radar (SAR) data in Han River Basin, Korea," *International Journal of Applied Earth Observation and Geoinformation*, vol. 83, 101930, 2019.

[4] T.-Y. Wu, A. Shao, and J.-S. Pan, "CTOA: Toward a Chaotic-Based Tumbleweed Optimization Algorithm," *Mathematics*, vol. 11, no. 10, 2339, 2023.

[5] T.-Y. Wu, H. Li, and S.-C. Chu, "CPPE: An Improved Phasmatodea Population Evolution Algorithm with Chaotic Maps," *Mathematics*, vol. 11, no. 9, 1977, 2023.

[6] F. Zhang, T.-Y. Wu, and Y. Wang, "Application of Quantum Genetic Optimization of LVQ Neural Network in Smart City Traffic Network Prediction," *IEEE Access*, vol. 8, pp. 104555-104564, 2022.

[7] K. Chen, R. Li, Y. Dou, Z. Liang, and Q. Lv, "Ranking Support Vector Machine with Kernel Approximation," *Computational Intelligence and Neuroscience*, vol. 2017, pp. 4629534-4629534, 2017.

[8] W. Zhou, Y. L. Sun, C. B. Xu, and G. Jian, "A Method for Discrimination of Ship Target and Azimuth Ambiguity in Multi-polarimetric SAR Imagery," *Journal of Radars*, vol. 4, no. 1, pp. 84-92, 2015.

[9] T. Zeng, D.-Y. Ao, C. Hu, and Z.-T. Cheng, "Multiangle Bistatic SAR Imaging and Fusion based on BeiDou-2Navigation Satellite System," *Journal of Radars*, vol. 3, no. 6, pp. 632-640, 2014.

[10] B. Peng, L. Zhang, and D. Zhang, "Automatic image segmentation by dynamic region merging," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3592-3605, 2011.

[11] Y. Xiang, F. Wang, L. Wan, and H. You, "An Advanced Multiscale Edge Detector Based on Gabor Filters for SAR Imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1522-1526, 2017.

[12] Z. Hu, Y. Wen, L. Liu, J. Jiang, R. Hong, M. Wang, and S. Yan, "Visual Classification of Furniture Styles," *ACM Transactions on Intelligent Systems and Technology*, vol. 8, no. 5, pp. 1-20, 2017.

[13] X.-R. Wang and Z.-C. Huang, "Painter Artistic Style Extraction Method Based on Color Features," *Journal of Computer Applications*, vol. 40, no. 6, pp. 1818-1823, 2020.

[14] S. Krinidis and V. Chatzis, "A Robust Fuzzy Local Information C-Means Clustering Algorithm," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1328-1337, 2010.

[15] F. Lamberti, A. Sanna, and G. Paravati, "Computer-assisted analysis of painting brushstrokes: digital image processing for unsupervised extraction of visible features from van Gogh's works," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, pp. 1-17, 2014.

[16] L. Du, L. Li, B. Wang, and J. Xiao, "Micro-doppler Feature Extraction Based on Time-frequency Spectrogram for Ground Moving Targets Classification with Low-resolution Radar," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3756-3763, 2016.

[17] L. Sun, S.-W. Ji, and J.-P. Ye, "Canonical Correlation Analysis for Multilabel Classification: A Least-Squares Formulation, Extensions, and Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 194-200, 2011.

[18] L. Jia, Y. Lei, E. Hendriks, and J. Z. Wang, "Rhythmic Brushstrokes Distinguish van Gogh from His Contemporaries: Findings via Automated Brushstroke Extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 6, pp. 1159-1176, 2012.

[19] X. Zhang, Y. Sun, G. Wang, Q. Guo, C. Zhang, and B. Chen, "Improved fuzzy clustering algorithm with non-local information for image segmentation," *Multimedia Tools and Applications*, vol. 76, no. 6, pp. 7869-7895, 2016.

[20] L. Kang, R.-S. Chen, N. Xiong, Y.-C. Chen, Y.-X. Hu, and C.-M. Chen, "Selecting Hyper-Parameters of Gaussian Process Regression Based on Non-Inertial Particle Swarm Optimization in Internet of Things," *IEEE Access*, vol. 7, pp. 59504-59513, 2019.

[21] C.-M. Chen, S. Lv, J. Ning, and J. M.-T. Wu, "A Genetic Algorithm for the Waitable Time-Varying Multi-Depot Green Vehicle Routing Problem," *Symmetry*, vol. 15, no. 1, 124, 2023.

[22] A. L. H. P. Shaik, M. K. Manoharan, A. K. Pani, R. R. Avala, and C.-M. Chen, "Gaussian Mutation–Spider Monkey Optimization (GM-SMO) Model for Remote Sensing Scene Classification," *Remote Sensing*, vol. 14, no. 24, 6279, 2022.
[23] W.-H. Hong, T. Lin, and Q.-M. Jia, "A Fuzzy Clustering Algorithm of Automatic Classification Based on EnFCM," *Advanced Materials Research*, vol. 989-994, pp. 1489-1492, 2014.
[24] A. Mekhmoukh and K. Mokrani, "Improved Fuzzy C-Means based Particle Swarm Optimization (PSO) initialization and outlier rejection with level set methods for MR brain image segmentation," *Computer Methods and Programs in Biomedicine*, vol. 122, no. 2, pp. 266-281, 2015.
[25] J.-J. Zhang, M. Fang, H. Wang, and X. Li, "Dependence maximization based label space dimension reduction for multi-label classification," *Engineering Applications of Artificial Intelligence*, vol. 45, pp. 453-463, 2015.
[26] Z. Khan and J. Yang, "Image segmentation via multi dimensional color transform and consensus based region merging," *Multimedia Tools and Applications*, vol. 78, no. 22, pp. 31347-31364, 2019.
[27] J. Nazarinezhad and M. Dehghani, "A contextual-based segmentation of compact PolSAR images using Markov Random Field (MRF) model," *International Journal of Remote Sensing*, vol. 40, no. 3, pp. 985-1010, 2018.
[28] D. M. Kumar, D. Satyanarayana, and M. N. G. Prasad, "An improved Gabor wavelet transform and rough K-means clustering algorithm for MRI brain tumor image segmentation," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 6939-6957, 2020.
[29] P. Mohapatra, S. Chakravarty, and P. K. Dash, "Microarray medical data classification using kernel ridge regression and modified cat swarm optimization based gene selection system," *Swarm and Evolutionary Computation*, vol. 28, pp. 144-160, 2016.