

Research Article

Bowen Hu and Yafei Yang*

Construction of a painting image classification model based on AI stroke feature extraction

<https://doi.org/10.1515/jisys-2024-0042>

received January 12, 2024; accepted February 22, 2024

Abstract: A large number of digital painting image resources cannot be directly converted into electronic form due to their differences in painting techniques and poor preservation of paintings. Moreover, the difficulty of extracting classification features can also lead to the consumption of human time and misclassification problems. The aim of this research is to address the challenges of converting various digital painting image resources into electronic form and the difficulties of accurately extracting classification features. The goal is to improve the usefulness and accuracy of painting image classification. Converting various digital painting image resources directly into electronic format and accurately extracting classification features are challenging due to differences in painting techniques and painting preservation, as well as the complexity of accurately extracting classification features. Overcoming these adjustments and improving the classification of painting features with the help of artificial intelligence (AI) techniques is crucial. The existing classification methods have good applications in different fields. But their research on painting classification is relatively limited. In order to better manage the painting system, advanced intelligent algorithms need to be introduced for corresponding work, such as feature recognition, image classification, etc. Through these studies, unlabeled classification of massive painting images can be carried out, while guiding future research directions. This study proposes an image classification model based on AI stroke features, which utilizes edge detection and grayscale image feature extraction to extract stroke features; and the convolutional neural network (CNN) and support vector machine are introduced into image classification, and an improved LeNet-5 CNN is proposed to achieve comprehensive assurance of image feature extraction. Considering the diversity of painting image features, the study combines color features with stroke features, and uses weighted K-means clustering algorithm to extract sample features. The experiment illustrates that the K-CNN hybrid model proposed in the study achieved an accuracy of 94.37% in extracting image information, which is higher than 78.24, 85.69, and 86.78% of C4.5, K-Nearest Neighbor (KNN), and Bi directional Long short-term Memory (BiLSTM) algorithms. In terms of image classification information recognition, the algorithms with better performance from good to poor are: the mixed model > BiLSTM > KNN > C4.5 model, with corresponding accuracy values of 0.938, 0.897, 0.872, and 0.851, respectively. And the number of fluctuation nodes in the mixed model is relatively small. And the sample search time is significantly shorter than other comparison algorithms, with a maximum recognition accuracy of 92.64% for the style, content, color, texture, and direction features of the image, which can effectively recognize the contrast and discrimination of the image. This method effectively provides a new technical means and research direction for digitizing image information.

Keywords: AI, stroke features, painting, image information, convolutional neural networks, K-means clustering algorithm, LeNet-5, support vector machine

* **Corresponding author: Yafei Yang**, School of Art, East China Jiaotong University, Nanchang 330031, China, e-mail: yangyafei202311@outlook.com

Bowen Hu: Architecture and Design College, Nanchang University, Nanchang 330031, China, e-mail: hubowen116@163.com

1 Introduction

With the boost of computer vision and artificial intelligence (AI) technology more and more automated painting methods have emerged. Machine learning and deep learning technologies are widely used in painting image classification, style transfer, and other fields, effectively optimizing traditional hand and digital painting methods. Painting, as a unique form of cultural expression, can effectively reflect the author's cognitive thinking about the world and emotional expression of real life. Countless excellent painting works carry the important crystallization of human civilization development and highlight the important spiritual power of social civilization development. With the boost of information technology and the emergence of mobile devices, digital painting forms have gradually emerged in people's lives, narrowing the distance between people and painting art appreciation. Convenient electronic devices can better help people understand and comprehend the artist and the spirit of the painting. The shape, length, width, density, and direction of strokes will outline different trajectory routes, and through the author's creation, different academic schools and styles of painting can be produced. It is also possible to learn and explore the potential laws in painting works. Therefore, strengthening the extraction of stroke features could markedly grasp the spatial correlation and orientation between painting works, thereby enhancing the accuracy of image classification and recognition. Meanwhile, the continuous development of AI technology has shown good results in image classification application scenarios. The AI algorithm has significant advantages in data information processing and image classification feature extraction. Therefore, the study proposes a painting image classification model based on AI stroke feature extraction, to provide better guidance for massive painting image processing.

The contribution of the article lies in the improved method, which enables the extraction of sample features and ensures the comprehensiveness of image feature extraction. It can effectively recognize the contrast and discrimination of images, and present image information by reflecting the contrast and correlation of images. At the same time, research methods can demonstrate good discrimination in classifying abstract, realistic, and romantic images. This method effectively provides a new technical means and research direction for digitizing image information. The article mainly includes the following parts. The first is the introduction, which is used to describe the background of the article. Next is the related work, which describes the current situation of the painting feature extraction and related image classification. The third Section is the design and implementation of painting image classification algorithms through stroke feature extraction under AI, image recognition and classification under deep learning methods, and consideration of multiple painting features. The fourth section is the results, which analyzed the painting classification situation, performance testing, and application analysis under this feature extraction method. Finally, the conclusion summarizes the content of the entire article.

2 Related works

Deep learning technology and related models have good algorithm performance in image classification and object detection, and related dataset construction and tool development are also emerging one after another. Multi branch and multi task deep neural networks provide technical tools for classifying painting styles and resolutions, and the sub region information presented by different resolution paintings on spatial transformation networks has significant feature differences [1,2]. Bianco et al. conducted different prediction task analysis on painting works covering multiple styles, and the accuracy of painting data classification based on genres exceeded 60% [3]. Based on the consistency of size requirements and sample imbalance in the classification and detection of abstract paintings, Bai and Guo constructed the dimensions of the feature map with local binary patterns. The neural network model containing AlexNet and spatial pyramid pool is fused at feature scales. The results indicate that this method can demonstrate high detection accuracy in dataset testing results [4]. Sandoval et al. used unsupervised adversarial clustering systems for automatic classification of art images, which involves designing label clustering modules and optimizing algorithms to connect unsupervised and supervised classification modules. The outcomes showcase that this method possesses better classification

performance relative to traditional clustering methods, and can effectively recognize application scenarios and image edge sharpness [5]. Jiang et al. proposed a new classification framework by combining discrete cosine transform with convolutional neural network (CNN) for the meticulous and freehand brushwork styles of Chinese ink painting, and achieved feature extraction of data information. The results indicate that this method has good classification performance [6]. Sun et al. combined CNN with genetic algorithms to design automatic network structure methods for image classification processing. The outcomes showcase that the fusion method possesses good classification accuracy and less resource consumption [7]. As an important feature of painting works, Sang and Kim have established a color emotion system in view of color combination technology and clustering algorithms, and applied it to the extraction of emotional features in painting. It can effectively convey the cultural value features contained behind the work [8]. Yang et al. transformed image grayscale features and used multi-scale grayscale co-occurrence matrix algorithm for extracting texture features of Chinese painting images, to better preserve image information. The experiment demonstrates that this method can effectively achieve multi feature differentiation in target recognition, reduce the interference of a large amount of irrelevant data, and have good generalization ability and target configuration recognition accuracy [9]. Wang and Huang used key areas to describe oil painting features, and used color features and naive Bayes classifier to classify image information. The outcomes showcase that this method possessed excellent classification accuracy in database experiments (all exceeding 90%), and it can effectively recognize the artist's artistic style [10].

Deep learning algorithms having high accuracy in image classification, selecting fewer features, and calculating image edges can ensure the integrity of their feature data [11]. CNN rarely involves modifications to fully connected layers in the architecture process, and most manually made image filters are prone to classification errors. Janke et al. used fully connected classifiers for testing the performance of different CNN classification algorithms, and believed that the introduction of classification algorithms such as support vector machine (SVM) and logistic regression can better achieve image feature classification processing than fully connected layers [12]. Bi et al. proposed feature learning based on genetic programming to extract and describe image classification information features. The outcomes showcase that this innovative learning algorithm can effectively filter feature descriptions and extract features from different classification tasks [13]. Klus and Gel used tensor networks to optimize the supervised learning problem and apply it to image classification. The results indicate that this method has better classification competitiveness compared to the basic neural network [14]. For hyperspectral images, Hu et al. used deep learning classification methods with fully connected features to classify wetland image information, and extract its spectral and texture features. The results indicate that this method exhibits good classification accuracy, and its balance of classification accuracy has been improved, with a classification accuracy difference of over 2% compared to other classification methods [15]. Zheng et al. focused on the classification problem of hyperspectral images by constructing a fast patchless global learning (FPGA) framework to achieve image framework construction. With the help of full convolutional networks and random layered sampling strategies, hyperspectral image classification and data information mining are achieved. And it ensures good convergence and extraction of global spatial information. The outcomes showcase that the designed FPGA framework exhibits good classification accuracy in the benchmark dataset [16]. The application of reinforcement learning is of great significance in computer vision problems, especially in image classification. Elizarov and Razinkov used deep neural networks for image classification and developed various reinforcement strategies to ensure the influence of experimental parameters on the results. The results indicate that the reinforcement learning method exhibits good applicability [17]. Li et al. proposed an end classification method in view of graphical neural networks, and extracted image features from multiple resolutions and scales, using saliency attention mechanism to update node features. This method has shown good feature extraction performance in Dongba painting image classification, with good accuracy and standard deviation [18]. In view of the advantages of deep learning in remote sensing image processing, Zhao et al. proposed image classification methods using stacked automatic encoding and deep residual networks, and applied them to hyperspectral remote sensing images. This means improving algorithm performance by reducing the dimensionality of feature data, constructing residual network modules, and processing data in batches. The dataset test results indicate that the classification method also exhibits good classification performance even in small sample sizes [19]. Liong proposed to use CNN for

feature description and weighted processing, and used normalized mutual information processing to sort features, and use SVM to classify. This embedded classification method exhibits strong data classification and robustness in experimental results [20]. For the common problem of information loss in feature extraction in view of local binary patterns, Zhao proposed image coding in view of color space distance to achieve feature extraction of color images. And it can effectively filter background information, with good applicability and effectiveness [21]. Iqbal optimize its efficiency by adjusting the deep learning optimizer and hyperparameter, and the algorithm spends less training time and occupies less memory resources [22]. Zhao proposed a color image LBP encoding method based on color space distance to address the issue of information loss in the application of local binary patterns in painting classification. The results indicate that the feature classification method exhibits good classification accuracy and application effectiveness [23].

Through the above research, it can be seen that deep learning technology and related models have good algorithm performance in image classification and object detection. These technologies can achieve high detection accuracy, feature extraction, classification performance, etc., in different fields. Most scholars use neural network algorithms, key region recognition, image label information, and color space distance to conduct research on image classification feature extraction and related fields. From this, it can be seen that these technologies have good application effects and are widely used. At the same time, some studies have also shown that deep learning technology has also been studied in the field of painting and has achieved relatively good results. These studies suggest that deep learning techniques can be used for feature extraction, image classification, and more in the field of painting. This has a positive promoting effect on the management of painting. But the existing methods mainly focus on processing the information features reflected in the image, and rarely analyze in view of the characteristics of the painting image itself. Neural networks exhibit good recognition accuracy in image classification. Therefore, this study utilizes the advantages of neural convolutional networks for image information classification and extracts painting image features in view of AI strokes to construct a classification model. This is to better achieve classification and recognition of image features.

3 Research on the construction of a painting image classification model based on AI stroke feature extraction

The study first proposes a classification method in view of AI stroke features, which filters and denoises painting images by processing their edges to extract feature information in terms of direction, texture, color, and other aspects. Subsequently, image information recognition and classification are implemented in view of CNN and SVMs, and in view of the construction of an image information library, painting images with multiple features are classified and processed. This is to improve the classification accuracy and recognition effect of painting images.

3.1 AI stroke feature extraction in painting image classification

As a form of brushstroke that expresses painting traces, the different shapes and directions of its combinations reflect the author's psychological activities and emotional expression during painting creation. Moreover, painters can exhibit different brushstroke characteristics through the comparison of different pigments, the weight and rhythm of their strokes, and the artistic style and emotional themes depicted also vary. Painting image classification is the classification and recognition of painting works according to different categories. Traditional painting image classification methods mainly rely on manual feature selection and classifier construction, requiring a large amount of professional knowledge and manual operation. And it has problems such as low recognition rate and poor robustness. While AI-based stroke features can extract and analyze features such as handwriting, contour, color, texture, etc., thereby achieving classification research on

painting images [24]. The research first utilizes edge detection methods to achieve feature extraction, and denoises and filters it to obtain stroke features. Edge detection is the feature extraction of the differences in edge line effects presented by operators, including abrupt edges, slow edges, and linear edges. The gray values of different image edges are different. The accuracy of common Laplace operator and Canny operator in edge feature detection is poor, and there are missing cases. Therefore, Sobel operator is introduced for detection. The Sobel operator mainly distinguishes features by judging the edge pixels of grayscale images, and uses convolutional kernels and algebraic weighting to obtain edge information during the feature extraction process. The mathematical expression of the Sobel operator is shown in equation (1).

$$\begin{cases} G_x = f_x(x, y) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \\ G_y = f_y(x, y) = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \end{cases} \quad (1)$$

where x represents the image, y represents the image, G_x is the convolution kernel of x and G_y is the convolutional kernel of y . When performing edge operations on a painting image, the principle of performing operations in view of one step gradient is that there is a gradient change difference in the grayscale values of pixels, and the corresponding change position of the gradient is the edge position of the region. Therefore, the grayscale value of pixel points can be expressed as equation (2).

$$G = \sqrt{G_x^2 + G_y^2}. \quad (2)$$

In the processing image information, wavelet threshold denoising algorithms are used to preserve the boundaries and local information of the image. However, the unreasonable threshold setting can easily lead to the extraction of image signals. Therefore, a new threshold function model is introduced to improve this problem, i.e., by converging the parts with larger absolute values in the coefficients; and high-order processing and zeroing are performed on those located at or below the threshold to ensure their preservation of information integrity. The improved algorithm represents the visible equation (3).

$$\widehat{W}_r = \begin{cases} \text{sign}(w) \left[|w| - 2T / \left\langle (2k+1) \left[1 + \exp\left(\frac{|w| - T}{2}\right) \right] \right\rangle \right], & |w| \geq T \\ \frac{k w^{2k+1}}{(2k+1) T^{2k}}, & T_0 \leq |w| < T \\ 0, & |w| < T_0 \end{cases} \quad (3)$$

where \widehat{W}_r is the wavelet coefficient after threshold processing; w is the original wavelet coefficient; $\text{sign}()$ is a symbolic function; T is the reference threshold; T_0 is the lower threshold; and k is the sample point. The selection of threshold values needs to be adjusted in view of certain parameters to achieve denoising. The Sobel operator can use a 3×3 filter to perform edge detection calculations on painted images converted to grayscale images, while considering the presence of interruptions or incomplete stroke shaping in the detected stroke features of the image. Therefore, to solve this problem, research is conducted using morphological operations to process edge detection to ensure the coverage of all stroke feature information. The calculation formula for defining morphological detection operators is shown in equation (4).

$$G = A \bullet B \circ B - A \circ B \bullet B. \quad (4)$$

where G represents the detection operator, \bullet represents closed operation, \circ represents open operations, B represents structural elements, and A represents images. Meanwhile, when extracting stroke features, it is necessary to filter and classify the selected stroke features to reduce unnecessary consumption of computational data caused by excessively sparse strokes. Therefore, it is necessary to perform image segmentation processing on it. First, region block segmentation is performed on the morphological processed image; subsequently, the gradient similarity of each sub drawing area is calculated using an elastic algorithm, and images

with similar similarity are merged and processed. It repeats this step until the division and sorting of the density size of each region are achieved, and the main indicator content of image recognition is in view of the first six stroke features with higher density values.

3.2 Painting image classification based on neural networks and SVMs

In response to the stroke features obtained from the research screening of painting classification, the study introduces CNN to abstract and automatically learn the features of painting images, and uses SVM to achieve image classification prediction. Neural networks, as a special type of AI technology, have a wide range of applications and good performance in the field of image classification, and image classification technology in view of AI algorithms has become one of the research hotspots in this field [25]. CNN, as an effective feedforward neural network to deal with computer image problems, is widely used in target detection, semantic segmentation, natural language processing, and other fields. It mainly extracts image features through multi-layer convolution and pooling operations, and inputs them into the fully connected layer for classification. During the training process, the algorithm continuously adjusts network weights through backpropagation, thereby improving classification accuracy. CNN is mainly composed of input layer, convolution layer, activation function, pooling layer, and full connection layer. It extracts information from the input data and compresses the features, and the output tensor data will have its own loss iteration in the role of the activation function, thus realizing the classification and recognition of information data. To solve the problem of gradient disappearance during training of deep network, the research designs T-ReLU activation function, whose mathematical expression is shown in equation (5).

$$f(l) = \begin{cases} l, & \text{if } l \geq 0 \\ a \left(\frac{1 - e^{-2l}}{1 + e^{-2l}} \right), & \text{if } l < 0. \end{cases} \quad (5)$$

where a is an adjustable parameter with an initial value of 0.25, and l is the sample data. The activation function has the effectiveness of Tanh function and ReLU function, which effectively avoids the output offset problem. The image after edge detection and denoising segmentation is input into a neural network with a size of 64×64 . The architecture of the five layer ConvNet model is shown in Figure 1.

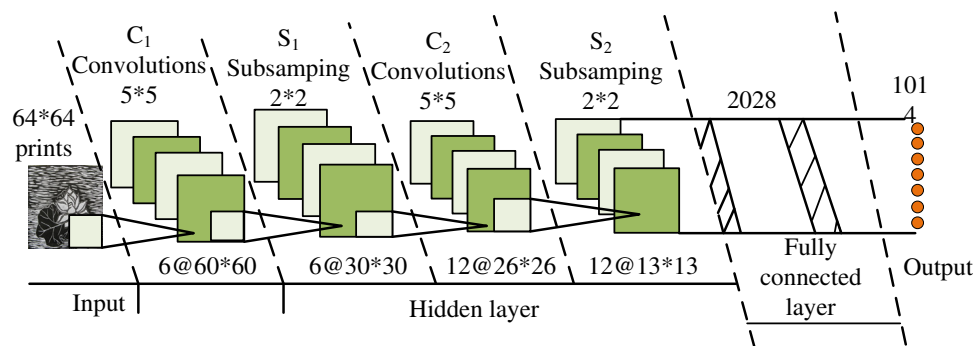


Figure 1: CNN network architecture of the five-layer convnet model.

In the experiment, the edge detected image is subjected to morphological operations, segmentation, and extraction, and then the image size of 64×64 is used as the input for CNN learning and training. In the C1 layer of CNN, the first convolutional layer filters the input data with six 5×5 convolutions to generate six 60×60 maps, learning some edge features from the original image pixels. In the S1 layer, with a subsampling rate of 2,

each map reduces the feature size to 30×30 by executing a maximum pooling layer, helping to extract significant edge features while reducing model parameters. In the C2 layer, the second convolutional layer filters data with 12 kernels of size 5×5 , generating 12 maps of size 26×26 , detecting simple shape features from edge features. On the S2 layer, 12 images reduce the feature size to 13×13 . In the S2 layer, the 2,028 dimensional vectors are obtained by designing a fully connected layer. Finally, 1,014 dimensional features are obtained in the output layer. Meanwhile, different receptors are reflected in corresponding areas after being stimulated, which is called Receptive field. The feature image pixels in CNN can effectively reflect the visual region situation. Figure 2 is a schematic diagram of the output results of the feature map.

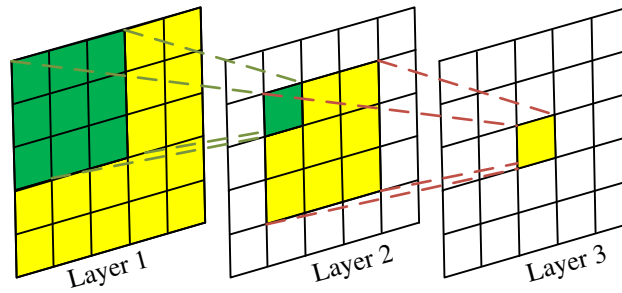


Figure 2: Output results of feature map.

Figure 2 shows the output result of the feature map, where each grid is equivalent to an element. The 3×3 green grid in Layer 1 represents the convolution kernel, and Layer 2 is the feature map output by the 3×3 convolution in Layer 1, with a feature size of 3×3 . And the green grid in Layer 2 is determined by the green grid in Layer 1, and its receptive field is the green grid area in Layer 1. Layer 3 is a feature map output by 3×3 convolution in Layer 2, with a feature size of 1×1 . The convolutional kernel function shown in this graph is consistent with the proposed five-layer ConvNet model, and the addition of convolutional kernels can enhance the nonlinear ability of the network. Therefore, an improved LeNet-5 CNN is proposed in view of the CNN network architecture of the five-layer ConvNet model. The network structure is shown in Figure 3.

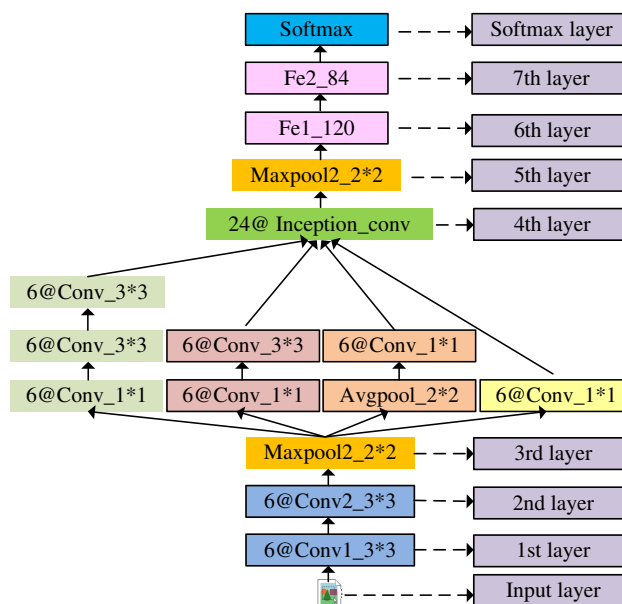


Figure 3: Schematic diagram of improved LeNet-5 CNN structure.

In the experiment, improvements were mainly made to the LeNet-5 CNN from the following three aspects. First, the average pooling in the LeNet-5 network was modified to maximum pooling, which could better extract image features. Second, using the convolutional kernel decomposition strategy, the 5×5 convolutional kernel in the first convolutional layer was replaced with two 3×3 convolutional kernels while maintaining the same receptive field. Finally, the convolutional module group inception was used to replace the third convolutional layer of the LeNet-5 network, making the network wider and deeper, and able to extract more image features. Subsequently, this study utilized SVMs for image classification. The core idea of SVM is to achieve the mapping of low dimensional data in high-order space, thereby improving the discrimination of the data. And it can distinguish different classified data by optimizing the optimal hyperplane, which has better accuracy and generalization ability. SVM classifies the features extracted from AI stroke features and constructs a painting image classification model. The linear transformation of SVM for classification problems can achieve optimal solutions or deficiencies. Figure 4 is a schematic diagram of the SVM model.

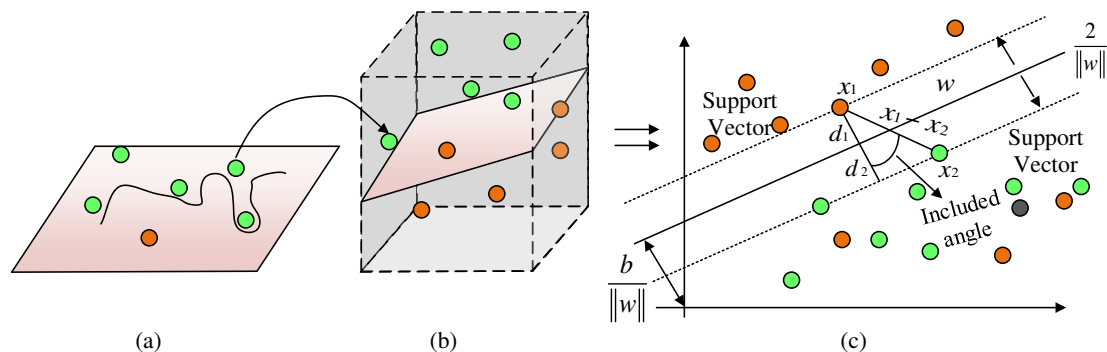


Figure 4: Schematic diagram of SVM model. (a) Input space, (b) feature space, (c) SVM.

Figure 4 shows a schematic diagram of the principle of the SVM model. The essence of SVMs is to transform non-linear classification problems into linear problems and obtain optimal solutions. In Figure 4, the low dimensional image (a) in the dataset is mapped to a high-dimensional feature space (b), making it linearly separable and maximizing the distance from the hyperplane, resulting in global optimization for the learner (c). When the SVM model classifies images, it mainly constructs the hyperplane between different databases in view of the decision function. Equation (6) is the mathematical expression of the decision function.

$$f(j) = \sum_{i=1}^n u_i k(j, J) + b. \quad (6)$$

where $f(j)$ represents the decision function; J is the training sample set; u is a generalized parameter; b is the deviation; and n represents the feature space dimension. And with the help of universal numerical values for image parameter values, the objective function of the kernel function is set as equation (7).

$$\min_{w, \varepsilon, b} \left\{ 0.5 \|w\|^2 + c \sum_{i=1}^n \varepsilon_i \right\}. \quad (7)$$

where ε is the relaxation variable, and c is the regularization parameter.

3.3 Painting image classification processing under multiple feature processing

Common painting images exhibit significant visual senses and color impact, and the color features they exhibit can effectively achieve classification of drawing images. Therefore, this study is in view of AI stroke features, taking into account their relationship with color features, to classify image datasets. Due to the influence of

their cultural backgrounds, there are significant differences in painting styles and strokes between China and the West. Research has been conducted to input color features of images into SVM classifiers for color vector, texture features, and other recognition. First, it transforms the image into a spatial coordinate system, i.e., RGB images are converted into HSV models. Figure 5 is a schematic diagram of the corresponding model.

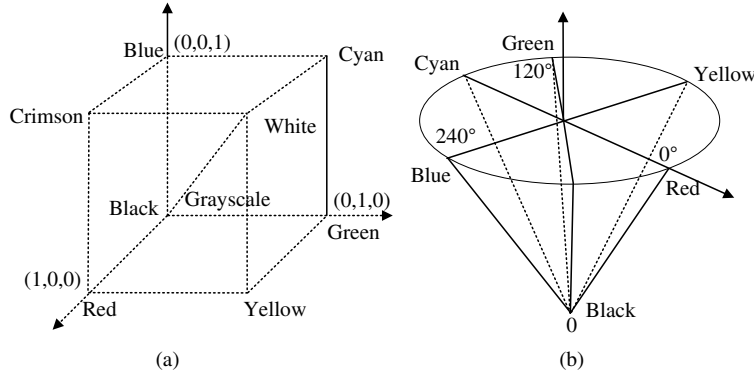


Figure 5: Schematic diagram of two models. (a) RGB color space model, and (b) HSV color space model.

The HSV model is different from the RGB color model which takes the red, green, and blue primary colors as the spatial axis. It reflects the color relationship of the image by taking hue, saturation, and lightness as parameters. The conversion formula for the two color models is shown in equation (8).

$$\begin{aligned}
 S &= [(V - N)/V] \quad \text{if } V \neq 0 \\
 H &= \begin{cases} \frac{60 \times (G - B)}{V - N} & \text{if } V = R \\ 120 + \frac{60 \times (B - R)}{V - N} & \text{if } V = G \\ 240 + \frac{60 \times (R - G)}{V - N} & \text{if } V = B \end{cases} \\
 V &= \max(R, G, B).
 \end{aligned} \tag{8}$$

where R denotes the red coordinates in the RGB model, G denotes the green coordinates in the RGB model, and B denotes the blue coordinate in the RGB model. H is the hue parameter in the HSV model, S is the saturation in the HSV model, and V is the lightness in the HSV model; N is the minimum value of spatial coordinates, $(\min(R, G, B))$. Subsequently, K-means clustering was performed on the HSV model to quantify its color space information. The K-means clustering algorithm, as an iterative clustering analysis algorithm, mainly classifies unlabeled sample data according to their attribute differences and ensures that the clustering results show small intra group differences. The similarity of sample distance is generally calculated using Euclidean distance, and its mathematical expression is shown in equation (9).

$$D(x_i, x_j) = \sqrt{\sum_{p=1}^q (x_{ip} - x_{jp})^2}. \tag{9}$$

where x_i represents the sample points for i ; x_j represents the sample points for j ; q represents the dimension of the data; and p is the total number of samples. The common clustering algorithm process is shown in Figure 6.

However, traditional clustering algorithms are sensitive to data noise and incomplete samples, which limits their processing efficiency. Therefore, in view of K-means clustering, the research enters information entropy for weighting to calculate the importance of index weights and then extract useful information. It standardizes the sample data and calculates the specific gravity and information entropy of the results, so the weighted Euclidean distance formula can be expressed as equation (10).

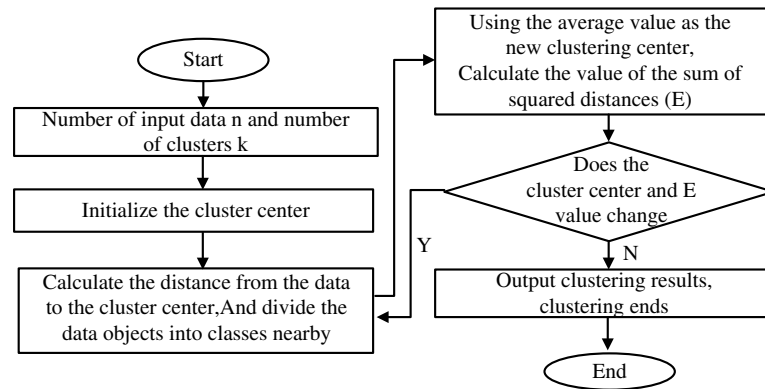


Figure 6: K-Means clustering algorithm flowchart.

$$D(x_i, x_j)' = \sqrt{\sum_{p=1}^q w_p (x_{ip}, x_{jp})^2}. \quad (10)$$

where w_p represents the weight coefficient. It calculates the weight coefficients of each indicator in the dataset, and uses the selected data object as the initial point for entropy weighted mean clustering, using weighted Euclidean distance to calculate the sample class and class center. Finally, the classification parameters under the distance result are implemented until the optimal parameter classification effect is achieved.

4 Analysis of painting image classification results based on AI stroke feature extraction

It analyzes the classification results of painting images proposed in the study, first setting up an experimental environment, namely, Intel (R) Core (TM) i7-7700 CPU @ 3.60 GHz processor; a 64 bit computer operating system with a memory size of 8GB; its development environment is set as PyCharm Community Edition; The programming language and deep learning framework are Python and Keras. The backend engine is Tensor Flow. In network training, the network parameters are initialized with default settings, with a learning rate of 0.001. Small batch input data are used, which are set to 128. Meanwhile, its design and painting database collect Chinese and western painting works. The painting content included in this sample set comprise characters, landscapes, and animals. It divides the collected painting dataset into training and testing sets

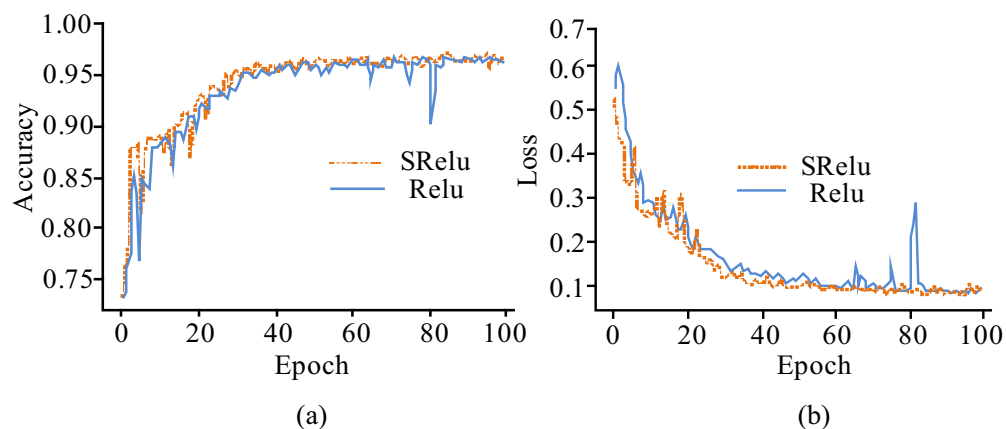


Figure 7: Image feature extraction results under two activation functions. (a) Accuracy curves and (b) loss rate curves.

in a 3:1 ratio. The performance analysis and evaluation of the painting image classification model tested in the research are shown in Figure 7.

The results in Figure 7 indicate that in terms of accuracy results, the accuracy curve of the original ReLU function is generally smaller than the improved SReLU function proposed in the study, with a large number of changing nodes. And there were significant fluctuations in the training batches at (0, 10) and 80 times, with an average accuracy of 86.74%. The accuracy change curve of the SReLU function proposed in the study gradually converges after training batches greater than 40 times. The overall average accuracy exceeds 90%. The maximum classification accuracy reached 96.24%. And the loss results indicate that the SReLU function curve is always below the ReLU function, and the data sample loss is relatively small. This is because SReLU function can fuse the right linear region of the ReLU function with the left soft saturation region of the Sigmoid function, and adjust the size of the left soft saturation region. Therefore, on the basis of ensuring convergence speed, the ReLU function can solve the problem of neuronal death. The study introduces adjustable parameters to improve the activation function of neural networks. To further analyze this improvement idea, it is compared with PSoftplus functions based on parameter correction [26], ELU functions [27], etc. The results are shown in Figure 8.

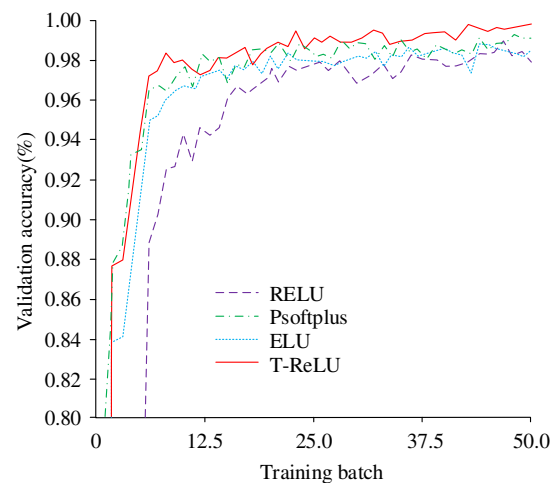


Figure 8: Verification accuracy of different functions.

The results in the graph indicate that under different training batches, the accuracy curves exhibited by different functions are different, and they all show an overall trend of increasing in the early stage and stabilizing in the later stage. The accuracy curve of ReLU function fluctuates significantly, and its effective accuracy in reaching convergence state is 97.2%. The average accuracy of the ELU function (98.12%) is higher than that of the PSoftplus function (97.65%). The accuracy curve of the PSoftplus function is similar to the accuracy curve of the improved T-RELU function proposed in the study, but there is still a slight fluctuation in the curve of the PSoftplus function after more than 25 training batches. The average accuracy of the improved function proposed in the study approaches 99.6%, indicating good effectiveness.

Subsequently, the improved CNN proposed in the study was analyzed for feature data extraction error results, as shown in Figure 9.

The results in Figure 8 indicate that there is a significant difference in data error results before and after the improvement of the CNN network. Specifically, in Figure 9(a), the predicted error value curve and the actual error value curve exhibited during the feature extraction of image data before the improvement of the CNN network have roughly the same trend. However, the fluctuation range of curve error under small sample sizes is 0.64%, and the overall error results have significant differences under different sample sizes. The characteristic curve in Figure 9(b) shows an error variation of 0.023%, and the overall fluctuation is relatively small due to changes in sample size. The reason for the discontinuity of the error situation in the results of

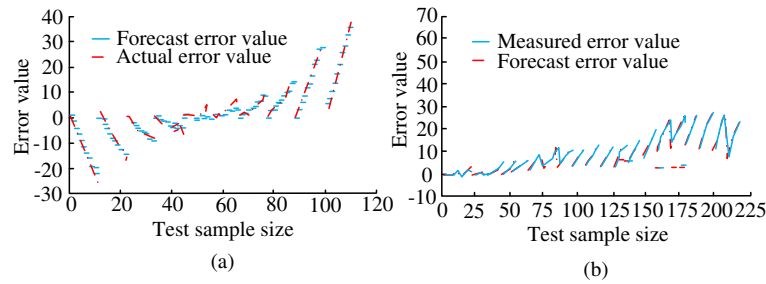


Figure 9: Data extraction error results (a) before and (b) after CNN network improvement.

Figure 8 is that the improved CNN network effectively reduces the variability of the feature input, but the removal of noise and outliers in the sample data cannot be completed to achieve 100% removal, and the variability between different types of image information also makes the amount of its features different. This is due to non-controllable objective factors caused by the subtle data deviation, and will not cause greater interference in the experimental results. Subsequently, the classification loss of the K-means clustering CNN proposed in the study was analyzed, and the results are shown in Figure 10.

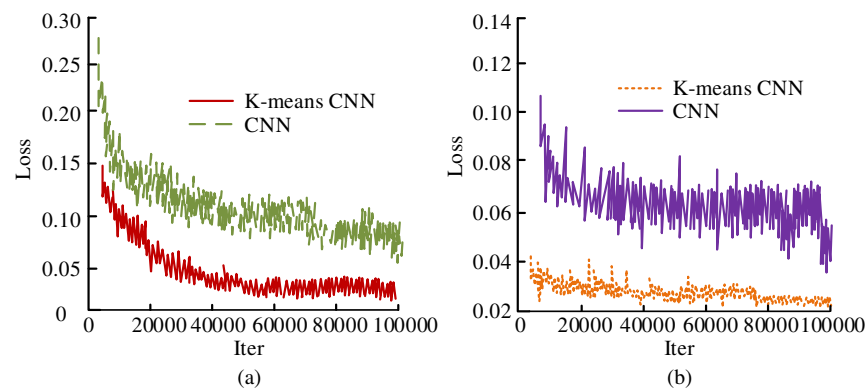


Figure 10: K Classification loss of mean clustering CNN. (a) Classification of loss and (b) return loss.

The results in Figure 10 indicate that the K-means clustering CNN algorithm exhibits lower classification loss and return loss than the CNN network. And it gradually tends to converge to a stationary state when the quantity of iterations exceeds 40,000 and 20,000, respectively, while the CNN network shows more obvious node fluctuations and losses. The reason for the discontinuity of the error feature curves in the results of Figure 10 may lie in the fact that the painting database constructed by the research has a large content of image information, and the type of data it covers makes its features differ to some extent. When analyzing the classification effectiveness of painting images, the differences in the amount of information caused by the objective factors of the images themselves cannot be completely ignored. Therefore, in the error analysis, the number of data category samples and the correlation between the sample features will make the extracted feature images have some discontinuous situations. Then, the accuracy effect of image information extraction using the classifier used in the study is analyzed, and the C4.5 Decision Tree, K-Nearest Neighbor (KNN), Bi directional Long Short-Term Memory (BiLSTM), etc., are selected for comparison. And the feature extraction effect was analyzed using the Precision Recall (PR) curve, as shown in Figure 11.

The results in Figure 11 indicate that the PR curves of the hybrid model and BLSTM model are closer to the lower right corner, with the accuracy of C4.5, KNN, and BiLSTM algorithms being 78.24, 85.69, and 86.78%, respectively. The K-CNN hybrid model proposed in the study achieved an accuracy of 94.37% in extracting image information. This is because the research has weighted the information entropy based on K-means

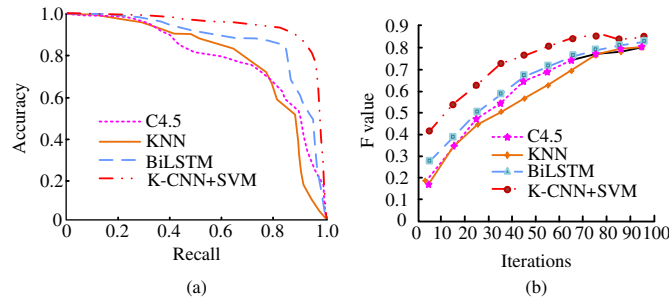


Figure 11: Comparison of PR curves and operating speeds of different models. (a) Comparison of PR changes under different models and (b) the number of iterations and F-value changes of different models.

clustering to achieve the calculation of the importance of indicator weights and the extraction of useful information. The weight coefficients of each indicator were calculated in the dataset, and the selected data object was used as the initial point for entropy weighted mean clustering. In the experiment, weighted Euclidean distance was used to calculate sample classes and class centers, ultimately improving the classification performance. In Figure 11(b), from the perspective of running speed, the hybrid model can effectively search for data samples, and its sample point fluctuation is relatively small. Its overall $F1$ value on the dataset is 85.16. Subsequently, the image feature recognition of different algorithms was analyzed, and the results are shown in Figure 12.

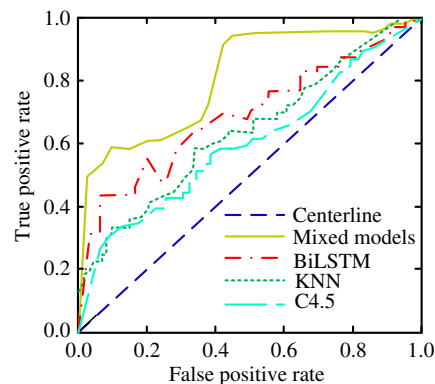


Figure 12: Classification performance of different classifiers.

The results in Figure 12 indicate that the algorithms with good performance in image classification information recognition have a difference ratio from good to poor: mixed model > BiLSTM model > KNN model > C4.5 model. The corresponding accuracy values are 0.938, 0.897, 0.872, and 0.851, respectively, and the number of volatility nodes in the mixed model are relatively small. This is because this network structure utilizes maximum pooling to extract image features, decompose the convolutional kernel, and replace a single third layer convolutional layer with a convolutional module group. This to some extent deepens the accuracy and comprehensiveness of image information extraction. At the same time, SVM is combined in the study for image classification. This method achieves the differentiation of different classification data through the optimization of the optimal hyperplane, and has good accuracy and generalization ability. Subsequently, the time consumption of four algorithm models for data classification was analyzed, and the results are shown in Figure 13.

The results in Figure 13 indicate that the KNN model and C4.5 model have significant changes in sample search time, and the overall fluctuation is relatively obvious, with an average search time of 16.38 and 17.23 ms.

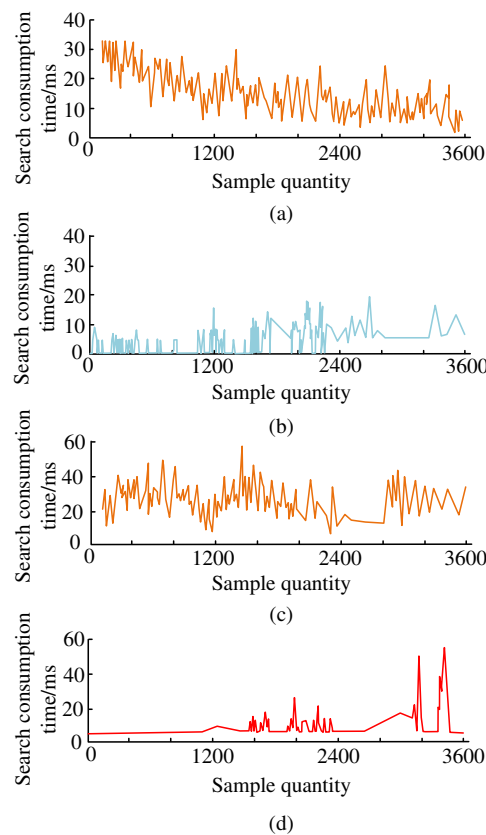


Figure 13: Time consumption of four algorithm models for data classification. (a) C4.5 model, (b) BiLSTM model, (c) KNN model, and (d) mixed model.

The BiLSTM model and the classification algorithm of the hybrid model proposed in the study are generally relatively stable in sample calculation; The average search time of the BiLSTM model is less than 15 ms, and it has good search performance in small sample sizes, but its fluctuation performance is poor compared to the mixed model. Subsequently, the recognition results of the image feature extraction method proposed in the study were analyzed, and the results are showcased in Figure 14.

Figure 14 shows the accuracy confusion matrix of image information classification in view of single color feature extraction and combination of color extraction and stroke feature extraction, where rows and columns represent real values and predicted values, respectively. Under the single color feature extraction in Figure 14(a), the algorithm achieves 75.38, 78.22, 92.16, 70.04, and 72.11% recognition accuracy for the style, content, color, texture, and direction of the image, with the maximum recognition accuracy being the recognition of image color. The classification algorithm in view of stroke features has a recognition accuracy of over 80% in all five types of image information, and the maximum and minimum recognition accuracies are 92.64 and 82.33%, respectively, which can effectively extract image information. Subsequently, the classification algorithms used in the study were subjected to image recognition of different genres and styles, and the spatial feature representation results are shown in Figure 15.

The results in Figure 15 indicate that the image classification algorithm in view of AI stroke features proposed in the study exhibits good discrimination in classifying abstract, realistic, and romantic images, and shows a clustered distribution between images of different styles. However, traditional image classification algorithms exhibit poor classification performance. And data analysis was conducted on the contrast and correlation of the classification image recognition, and the results are shown in Figure 16.

The results in Figure 16 indicate that the classification algorithm in view of AI stroke features can better reflect the contrast and correlation of images, with corresponding average values of 1.12 and 0.97, respectively,

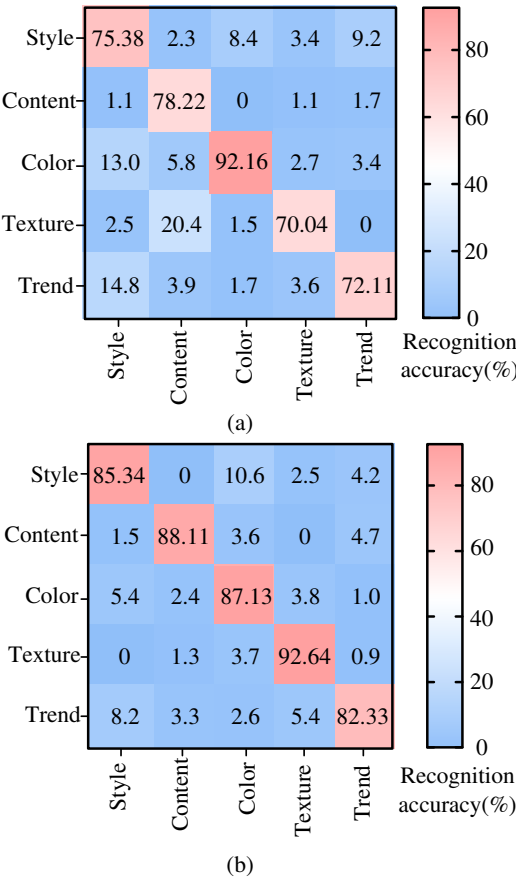


Figure 14: Accuracy of two image feature extraction confusion matrix. (a) Single color feature extraction and (b) color recognition+stroke feature recognition.

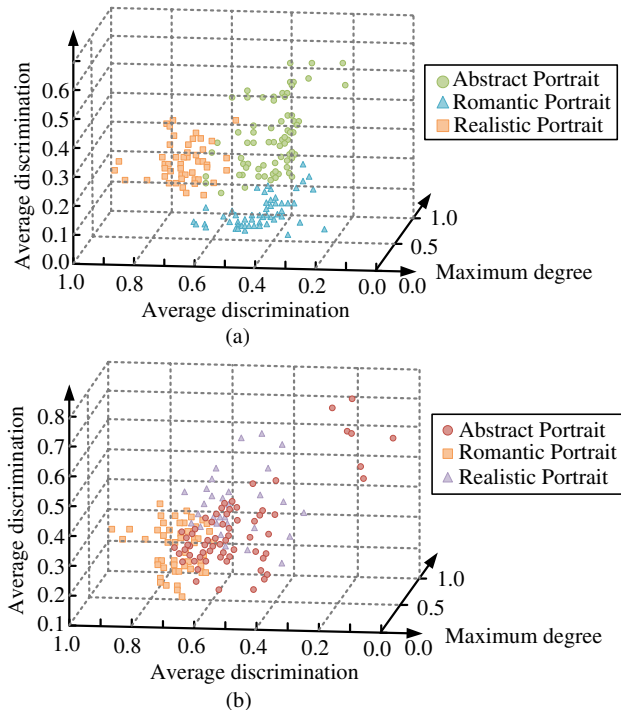


Figure 15: The feature space classification of samples. (a) Traditional image classification algorithms and (b) image classification algorithm based on AI stroke features.

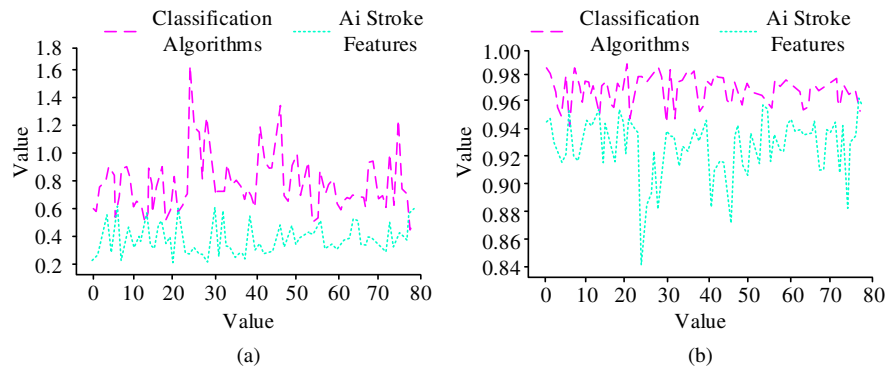


Figure 16: Contrast and correlation of different classification image recognition. (a) Contrast and (b) relevance.

which can better present image information. However, the contrast of traditional algorithms after image classification is relatively low, indicating that their ability to grasp the brightness relationship of images is not very good. The overall correlation changes greatly, with average contrast and correlation values of 0.45 and 0.91. The above outcomes demonstrate that the algorithm in view of AI stroke features has good image classification results and good recognition performance for information.

5 Conclusion

As an important part of image processing, the extraction of feature information in painting has important practical value. Research is conducted on image classification in view of AI stroke features, and simulation results are analyzed. The outcomes demonstrated that the accuracy change curve of the SReLU function proposed in the study gradually converged after more than 40 training batches, with an overall average accuracy of over 90% and a maximum classification accuracy of 96.24%. And the K-means clustering CNN algorithm gradually converges to a stable state when the number of iterations is greater than 40,000 and 20,000, respectively. The accuracy of image information extraction reaches 94.37%, and the fluctuation of sample search is small, with an $F1$ value of 85.16. In terms of resource consumption time, the average search times of KNN model and C4.5 model samples are 16.38 and 17.23 ms, while the average search time of BiLSTM model is less than 15 ms, and its performance is poor compared to the hybrid model. And the accuracy of the single color feature extraction algorithm for the style, content, color, texture, and direction of the image is 75.38, 78.22, 92.16, 70.04, and 72.11%. The classification algorithm in view of stroke features has a recognition accuracy of over 80% in all five types of image information, with a maximum recognition accuracy of 92.64% and a minimum recognition accuracy of 82.33%, respectively. This can effectively distinguish the contrast and correlation of image information, with corresponding average values of 1.12 and 0.97, which are much higher than traditional algorithms' 0.45 and 0.91. The algorithm in view of AI stroke features has good image classification results, and the quality and accuracy of information recognition are high. In the part of extracting stroke features, although detection algorithms and morphological operations can effectively detect and process image features, more comprehensive methods need to be proposed in the future to ensure comprehensiveness and completeness of stroke feature selection. At the same time, the content contained in painting images is relatively complex, including texture, color, and other contents in addition to strokes. Therefore, in the future, research can comprehensively consider multiple features and design classification models with multiple features. Similarly, paying attention to appropriate parameter tuning for image selection and improving the database's indexing and hardware capabilities are also important aspects that research needs to focus on in the future.

Funding information: The authors state no funding involved.

Author contributions: Bowen Hu wrote original draft; Yafei Yang conducted data analysis and method guidance.

Conflict of interest: The authors declare that they have no conflict of interest.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- [1] Oslund S, Washington C, So A, Chen T, Ji H. Multiview robust adversarial stickers for arbitrary objects in the physical world. *J Comput Cogn Eng*. Nov. 2022;1(4):152–8. doi: 10.47852/bonviewJCE2202322.
- [2] Wu H, Liu Q, Liu X. A review on deep learning approaches to image classification and object segmentation. *Comput Mater Contin*. Jan. 2019;58(2):575–97. doi: 10.32604/cmc.2019.03595.
- [3] Bianco S, Mazzini D, Napoletano P, Schettini R. Multitask painting categorization by deep multibranch neural network. *Expert Syst Appl*. Jun. 2019;135(11):90–101. doi: 10.1016/j.eswa.2019.05.036 Get rights and content.
- [4] Bai R, Guo X. Automatic orientation detection of abstract painting. *Knowl-B Syst*. Jun. 2021;227(3):107240. doi: 10.1016/j.knosys.2021.107240.
- [5] Sandoval C, Pirogova E, Lech M. Adversarial learning approach to unsupervised labeling of fine art paintings. *IEEE Access*. Jun. 2021;9:81969–85. doi: 10.1109/ACCESS.2021.3086476.
- [6] Jiang W, Wang Z, Jin JS, Han Y, Sun M. DCT-CNN-based classification method for the Gongbi and Xieyi techniques of Chinese ink-wash paintings. *Neurocomputing*. Feb. 2019;330(12):280–6. doi: 10.1016/j.neucom.2018.11.003.
- [7] Sun Y, Xue B, Zhang M, Yen GG, Lv J. Automatically designing CNN architectures using the genetic algorithm for image classification. *IEEE Trans Cybern*. Sep. 2020;1(99):1–15. doi: 10.1109/TCYB.2020.2983860.
- [8] Sang HL, Kim JY. Classification of the era emotion reflected on the image using characteristics of color and color-based classification method. *Int J Softw Eng Knowl Eng*. Aug. 2019;29(8):1103–23. doi: 10.1142/S0218194019400114.
- [9] Yang D, Ye X, Guo B. Application of multitask joint sparse representation algorithm in chinese painting image classification. *Complex*. Jan. 2021;2:1–11. doi: 10.1155/2021/5546338.
- [10] Wang Q, Huang Z. Painter artistic style extraction method based on color features. *J Comput Appl*. Dec. 2019;40(6):1818–23. doi: 10.11772/j.issn.1001-9081.2019111886.
- [11] Jaruenpunyasak J, Duangsoithong R. Empirical analysis of feature reduction in deep learning and conventional methods for foot image classification. *IEEE Access*. Mar. 2021;1(99):53133–45. doi: 10.1109/ACCESS.2021.3069625.
- [12] Janke J, Castelli M, Popovic A. Analysis of the proficiency of fully connected neural networks in the process of classifying digital images. Benchmark of different classification algorithms on high-level image features from convolutional layers. *Expert Syst Appl*. Nov. 2019;135(11):12–38. doi: 10.1016/j.eswa.2019.05.058.
- [13] Bi Y, Xue B, Zhang M. An effective feature learning approach using genetic programming with image descriptors for image classification [Research Frontier]. *IEEE Comput Intell Mag*. Apr. 2020;15(2):65–77. doi: 10.1109/MCI.2020.2976186.
- [14] Klus S, Gel P. Tensor-based algorithms for image classification. *Algorithm*. Nov. 2019;12(11):240. doi: 10.3390/a12110240.
- [15] Hu Y, Zhang J, Ma Y, Li X, Sun Q, An J. Deep learning classification of coastal wetland hyperspectral image combined spectra and texture features: A case study of Huanghe (Yellow) River Estuary wetland. *Acta Oceanol Sin*. Sep. 2019;38(5):142–50. doi: 10.1007/s13131-019-1445-z.
- [16] Zheng Z, Zhong Y, Ma A, Zhang L. FPGA: Fast patch-free global learning framework for fully end-to-end hyperspectral image classification. *IEEE Trans Geosci Remot Sens*. Feb. 2020;99:1–15. doi: 10.1109/TGRS.2020.2967821.
- [17] Elizarov AA, Razinkov EV. Image classification using reinforcement learning. *Russ Digit Libr J*. May. 2020;23(6):1172–91. doi: 10.26907/1562-5419-2020-23-6-1172-1191.
- [18] Li K, Qian W, Wang C, Xu D. Dongba painting few-shot classification based on graph neural network. *J Comput Aid Des Comput Graph*. Nov. 2021;33(7):1073–83. doi: 10.3724/SP.J.1089.2021.18618.
- [19] Zhao J, Hu L, Dong Y, Huang L, Weng S, Zhang D. A combination method of stacked autoencoder and 3D deep residual network for hyperspectral image classification. *Int J Appl Earth Obs Geoinf*. Oct. 2021;102(4):102459. doi: 10.1016/j.jag.2021.102459.
- [20] Liong ST, Huang YC, Li S, Huang Z, Ma J, Gan YS. Automatic traditional Chinese painting classification: A benchmarking analysis. *Comput Intell*. May. 2020;36(3):1183–99. doi: 10.1111/coim.12328.
- [21] Zhao Q. Research on the application of local binary patterns based on color distance in image classification. *Multimed Tool Appl*. May. 2021;80(18):27279–98. doi: 10.1007/s11042-021-10996-9.
- [22] Iqbal I, Odesanmi GA, Wang J, Liu L. Comparative investigation of learning algorithms for image classification with small dataset. *Appl Artif Intell*. Jun. 2021;35(5):1–20. doi: 10.1080/08839514.2021.1922841.

- [23] Sheng J, Li Y. Classification of traditional Chinese paintings using a modified embedding algorithm. *J Electron Imag.* Mar. 2019;28(2):1–11. doi: 10.1117/1.JEI.28.2.023013.
- [24] Ren Z, Fang F, Yan N, Wu Y. State of the art in defect detection based on machine vision. *Int J Precis Eng Manuf-Green Technol.* May. 2022;9(2):661–91. doi: 10.1007/s40684-021-00343-6.
- [25] Zhang Y. Image feature extraction algorithm in big data environment. *J Intell Fuzzy Syst.* Oct. 2020;39(1):1–10. doi: 10.3233/JIFS-179996.
- [26] Salam A, Hibaoui AEI, Saif A. A comparison of activation functions in multilayer neural network for predicting the production and consumption of electricity power. *Int J Electr Comput.* February. 2021;11(1):163–70. doi: 10.11591/ijece.v11i1.pp163-170.
- [27] Yao W, Wang C, Sun Y, Zhou C. Robust multimode function synchronization of memristive neural networks with parameter perturbations and time-varying delays. *IEEE Trans Syst Man Cybern: Syst.* June. 2020;1(99):2–15. doi: 10.1109/TSMC.2020.2997930.