www.cad-journal.net
Computer-Aided Design and Applications

Analysis of Performance Characteristics of Folk Art Images Based on Deep Learning

Yulu Huo[1] and Peng Zhao[2]

[1]Department of Humanity and Law, Zhengzhou Technology and Business University, Zhengzhou, Henan 450000, China, 3000013913@ztbu.edu.cn
2Department of Humanity and Law, Zhengzhou Technology and Business University, Zhengzhou, Henan 450000, China, hnnhzp@163.com

Corresponding author: Yulu Huo, 3000013913@ztbu.edu.cn

Abstract. Folk art images record the growth of human traditional culture. The classification study of folk-art images is helpful to sort out painting resources, analyze the performance characteristics of folk-art images, and better inherit China traditional culture. The classification of folk-art images is of great significance to the management and use of art works. Traditional image detection methods need to manually extract image features, which is complicated and requires professional knowledge, making it impossible to fully extract image feature information. Convolutional neural network (CNN) can extract rich and effective image features, so the feature expression method based on learning has better performance. A linear and nonlinear function is proposed as the activation function. The research results indicate that the deep learning algorithm tested in this article can perform accurate image recognition in a short period of time, which is superior to the recommendation performance of traditional algorithms and can play an obvious advantage in CAD image feature recognition of folk-art images.

Keywords: Art Images; Deep Learning; CAD; Trick Recognition
DOI: https://doi.org/10.14733/cadaps.2024.S1.14-30

1 INTRODUCTION

Painting has played a very important role in the growth of human civilization. It has transformed people's understanding and feelings about the world into visual images. It is an important way for people to know and express the world through beautiful works of art. Beautiful works of art are spiritual food for modern people to cultivate their sentiments and cultivate their self-cultivation, which can help people release the pressure of life and work and make them happy to work and live. Therefore, it is of great significance to study the image expression of folk art. The classification of folk-art images is of great significance to the management and use of art works. Traditional image detection methods need to manually extract image features, which is complicated and requires professional knowledge, making it impossible to fully extract image feature information. Traditional image features that have undergone structured processing can perform excellent information image feature analysis in image data conversion. It analyzes the training of structured image feature model classifiers.

Folk art is an ancient art form, which is the crystallization of the wisdom of the people, with profound cultural connotation and broad social foundation. With the development of deep learning technology, we can use computer vision and Natural language processing and other technologies to analyze the characteristics of Folk art images, so as to better protect and inherit Folk art. First, we can use computer vision technology to analyze the image of Folk art. Specifically, we can digitize the image of Folk art and transform it into digital images that can be processed by computers. Then, we can use image processing algorithms to process these digital images and extract feature information from them. For example, we can use Convolutional neural network (CNN) to extract and analyze the patterns, colors, textures and other features of Folk art images. In this way, we can further understand the manifestation and characteristics of Folk art images, and provide more scientific and effective support for the protection and inheritance of Folk art. Secondly, we can also use Natural language processing technology to analyze the description and interpretation of Folk art images. Specifically, we can transform the description and interpretation of Folk art images into text forms, and then use Natural language processing technology to process them, such as word segmentation, part of speech tagging, Named-entity recognition, etc. In this way, we can better understand the cultural connotation and social background of the image of Folk art, and provide more comprehensive and systematic support for the protection and inheritance of Folk art. Finally, we can also combine computer vision and Natural language processing technology to comprehensively analyze the image of Folk art. Specifically, we can combine digital images and text descriptions of Folk art images, and use multimodal fusion technology to combine the two. Then, we can extract and analyze the characteristics of Folk art through the in-depth learning model, so as to more comprehensively understand the manifestation and cultural connotation of the image of folk art.

In conclusion, deep learning technology can provide strong support for the analysis of Folk art image performance characteristics. Through computer vision, Natural language processing and other technologies, we can more deeply understand the manifestation and cultural connotation of Folk art images, and provide more scientific, comprehensive and systematic support for the protection and inheritance of Folk art. Folk art images record the growth of human traditional culture. The classification study of folk-art images is helpful to sort out painting resources, analyze the performance characteristics of folk-art images, and better inherit China traditional culture. Art image detection generally refers to the use of computer technology to predict the content, painting style and age of art works. The digitized art images of art works are input into the computer, which, with its powerful calculation and analysis capabilities, can represent the image features and the overall structure of the image. Art works are of great significance in the study of human history, science, art and culture, but it is difficult to obtain art works, which makes the research work of researchers more difficult. CNN can extract rich and effective image features, so the feature expression method based on learning has better performance. The model parameters of natural image training and the model parameters of art image training have great sharing. In this paper, the Simulated Annealing Algorithm (SAA) is applied to CNN for art image recognition, and the influence of activation function on the accuracy of image detection is further studied. A linear and nonlinear function is proposed as the activation function, and a training network is established to form a classifier. Then, folk art CAD images with different artistic attributes are classified, and then the performance characteristics of folk-art images are analyzed.

CAD (Computer-aided design) and deep learning technology are widely used in the image representation of Folk art. CAD technology can help designers and artists design and create more efficiently, while deep learning technology can help computers automatically learn and generate artworks. CAD technology can be used to design and produce prototypes, models and finished products of folk art. For example, using CAD software can design digital models of traditional wood carving, thereby more accurately controlling the carving process. CAD technology can also be used to produce traditional handicrafts such as ceramics, fabrics, and embroidery. By using CAD software, designers can create complex patterns and designs and transform them into actual folk art. Deep learning technology can be used to generate and optimize works of art. For example, using the Generative adversarial network (GAN) can generate realistic Folk art images, such as wood carving, ceramics and fabrics. Deep learning technology can also be used to optimize the quality and style of artistic works. For example, using depth learning technology can transform the style of traditional hand drawn images to make them look like they are drawn by folk artists. In general, CAD and deep learning technology are widely used in the image representation of Folk art, which can help designers and artists design and create more efficiently, and can also help computers automatically learn and generate art works.

Due to the differences of human living environment, ethnic inheritance and other factors, in the long-term relatively stable natural and cultural environment, different national color forms and color views have gradually formed. In the process of image detection, because image features contain various visual information such as pixels, colors, shapes and textures, when classifying images, different classification information is emphasized, and image feature selection is also different, and the final result will be different because of the angle selection. Folk art is dominated by freehand brushwork, and there is a big difference between images and objects, and the number of images is limited, which is not conducive to image detection. In this paper, an analysis method of performance characteristics of folk-art images based on CAD and deep learning is proposed, and its main contributions are as follows:

(1) This paper studies the application of image feature analysis to model and analyze the feature analysis of Folk art images.

(2) Aiming at the problem of image feature representation in folk art image detection and image retrieval, this model proposes an image representation method based on sparse representation and a deep learning algorithm based on hierarchical feature extraction to analyze the performance characteristics of folk-art images.

In this paper, CNN model of feature recognition of art images is constructed, and its convolution operation process is explained, and then experiments are carried out on the general data set of folk art. The test results show that the deep learning algorithm in this study can realize accurate image feature recognition in a short time.

2 RELATE WORK

To use a CAD image mapping dataset for predicting different network parameters, the following steps can be followed: first, the CAD image mapping dataset needs to be collected and prepared. This dataset should contain CAD images and corresponding mapping data under different network parameters. For each data point, it is necessary to ensure that the required features and labels are included. Divide the dataset into training and validation sets. The training set is used to train the model, while the validation set is used to adjust model parameters and evaluate model performance. Select an appropriate deep learning model, such as Convolutional neural network (CNN) or Recurrent neural network (RNN). Build models using Python and appropriate deep learning frameworks such as TensorFlow or PyTorch. The training set is input into the model, and the Backpropagation is used to calculate the gradient and optimize the model parameters. During the training process, validation sets can be used to monitor the performance of the model and make adjustments as needed.

Evaluation Model: Use a test set to evaluate the performance of the model. Calculate indicators such as accuracy, recall, and F1 score of the model. Adjust and optimize based on the evaluation results, and use the trained model to predict new CAD image mapping data. Input new data into the model and output prediction results. Analyze the prediction results to determine the performance and accuracy of the model. Further optimization and improvement as needed. It should be noted that when making predictions, it is necessary to ensure that the model has been fully trained and adjusted to the optimal parameters. In addition, appropriate preprocessing and feature extraction of the data are also necessary to ensure the performance and accuracy of the model.

Bappy et al. [1] conducted a high confidence image recognition analysis using Berning Magi manipulation mapping. By discriminating the regions of the mapping dataset of CAD images, a prediction encoding architecture for different network parameters was constructed. After strict validation with different datasets, the proposed method can achieve high-precision image operations. Bayarri et al. [2] analyzed the effects of changes in organic elements based on natural pigments. Its hyperspectral analysis of visible and near-infrared light with CAD from various countries has been meticulously carried out through the recognition of graphic forms and the preservation of superimposed states. By establishing a visual enhancement of color materials, the preprocessing of graphics was continued. Benzon et al. [3] performed periodic representation of digital virtual structures in drone images. By manufacturing 3D geometric reconstruction of drone images, it analyzed the geometric representation of digital twins in images. Capotorto et al. [4] analyzed the perspective analysis of painting space under the 3D model of artistic landscape composition CAD. It adopts the method of linear perspective to reconstruct the three-dimensional model of landscape composition through geometric rules of perspective works. Ding et al. [5] devoted themselves to the study of ancient art images by utilizing the information modeling of CAD architectural images. The results indicate that the proposed architectural information images are different from other ancient buildings in terms of internal significance. Gu et al. [6] conducted image object detection of remote sensing objects and analyzed the detection scene targets under different educational objectives. Through reviewing the technology of sustainable development depth learning method, it analyzes the task recognition of scene image recognition category. Hong et al. [7] analyzed the spatial spectral sampling model capability. A feature extraction method for CAD single models was developed through in-depth research on regular image data of convolutional networks. The results show that Convolutional neural network has advantages over single model in image fusion strategy. Hong et al. [8] analyzed the digital representation of scene style CAD in Chinese classical private gardens. Through the transformation of the aesthetic style of Shan shui based on neural network, it conducted in-depth learning of virtual scene style painting analysis. The network verified the stylized expression of landscape Shan shui by predicting the multi content scene style transformation of the network style. Jaspe et al. [9] introduced an image shape material data model based on CAD structural informatization. By creating different data tools, it analyzed the image editing tool dataset of multi-layer datasets. By embedding interactive visualization to enhance analysis, it constructs an associated architecture for layer selection. McCartney et al. [10] conducted a modernist aesthetic analysis. Through the design analysis of visual art in Computer-aided design, it reflects the irreplaceable role of computer-based modernism aesthetics in the analysis of artistic image performance characteristics. Murthy et al. [11] analyzed the algorithm application of computer Assistive technology in the field of image vision. It provides a detailed introduction and analysis of the indicator evaluation of deep learning technology in object detection. Nordin et al. [12] used CAD's graphical binary method for rapid detection of high-resolution images with reciprocal horizontal thresholds. It analyzed the defect detection and repair process of image automatic grinding tools, and constructed the defect repair process of automatic molds by analyzing the high-resolution scanning results of dyed drawings. Santoso et al. [13] conducted a type construction of CAD image art. The development of type technology was analyzed through the overall recognition of Wayang image objects. Xu and Jiang [14] have conducted high-quality talent cultivation in the field of artificial intelligence art design. A computer-based course system optimization analysis was conducted on the efficient art design of machine learning. Improved the knowledge environment for art teaching under artificial intelligence. Xue [15] designed a micro pattern mural system based on a CAD assisted system. By simulating the exchange method of visual aesthetic information conveyed by mural patterns, an alternating vertical mapping of wall painting patterns was achieved, which has a strong aesthetic effect. Zhao et al. [16] conducted computer-aided painting rendering art creation. Currently, digital painting is becoming a technological breakthrough in the professional painting industry. With the help of intelligent computer technology, it has carried out intelligent technology for artistic vision, image geometry rendering. This technology not only preserves traditional image formats in digital processing of image information, but also satisfies traditional image optimization.

3 METHODOLOGY

3.1 CNN Model for Feature Recognition of Art Images

The feature extraction and recognition of large-scale art image databases can be completed through the following steps. Convert each image into a feature vector, commonly used methods include color histograms, color sets, color moments, color aggregation vectors, color correlation maps, etc. These methods have their own advantages and disadvantages. For example, color histograms are not affected by image rotation and translation changes, and further normalization can avoid the influence of image scale changes. However, they do not express information about color spatial distribution. Feature recognition involves inputting the extracted feature vectors into a classifier, such as support vector machines, naive Bayes, decision trees, etc., to complete the task of recognizing artistic images. Convolutional neural network (CNN) has certain advantages in feature recognition of artistic images. The characteristics of Convolutional neural network include: Convolutional neural network can effectively process image recognition tasks with translation, scaling and Rotational invariance, which are very important for the recognition and analysis of artistic images. The Convolutional neural network can directly act on the input samples, train the network through the samples and finally realize the detection task. Convolutional neural network also has excellent performance in face detection. Convolutional neural network can accurately recognize and detect faces with any size, position, pose, direction, skin color, facial expression and lighting conditions in the image. In the aspect of character recognition, Convolutional neural network can effectively extract the features of characters, carry out correlation analysis, find the features that can best represent characters, and remove the features that are irrelevant to classification and autocorrelation. However, Convolutional neural network still has some challenges and limitations, such as feature extraction and recognition for large-scale art image databases, which may require a lot of computing resources and time. In addition, Convolutional neural network needs a large number of training data and labels to train, which may lead to high costs.

Image representation has a lot to do with the characteristics of the image itself, that is, image representation is closely related to its specific application fields. In the early days, image processing and machine learning methods could not meet the requirement of distinguishing features to a certain extent, so image feature modeling based on human prior knowledge and subjective creativity occupied an extremely important position. The basic model of CNN based on deep learning is shown in Figure 1.

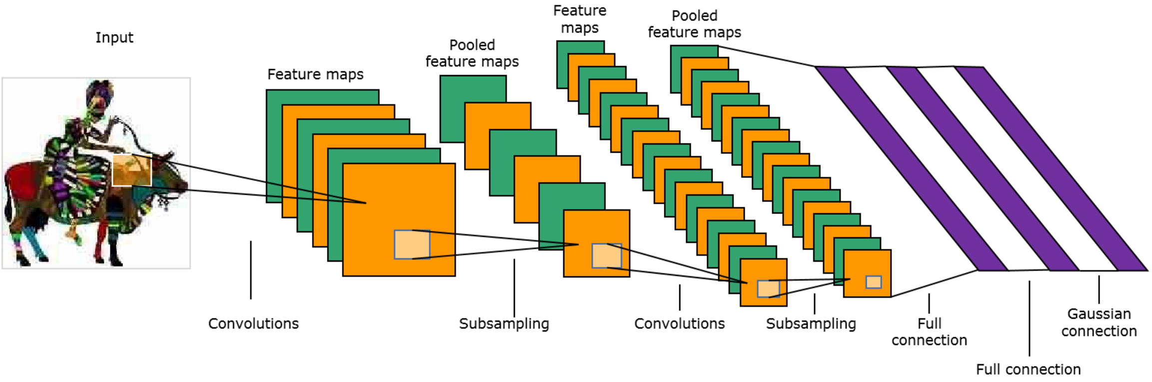

Figure 1: Basic model of CNN.

Convolution layer and down-sampling layer are often combined to form a convolution group. The main function of convolution layer is feature detection, promotes the general method of deep CNN model construction, and multiple convolution groups are alternately arranged to realize feature extraction layer by layer.

Based on the characteristics of local receptive field, CNN can focus on learning the image content of a specific area, which also makes CNN more adaptable to the displacement of objects in the image. Because the convolution kernel is locally connected with the feature map, and the weight parameters of the convolution kernel can be shared in the calculation of different feature maps, the calculation cost of CNN is reduced. In the research of pattern recognition, features are the description and representation of the things studied. We always hope to extract the effective feature expression for the recognition task, so that the computer can have better results in the subsequent recognition task. Especially for the visual field that this paper focuses on, features are an important part of determining the quality of the whole system.

Feature extraction is the key step of folk-art image detection. The key link of Folk art image convolution network detection includes data preprocessing. Including image acquisition, image annotation, image enhancement, etc., used to increase the amount of training data and improve classification accuracy. Construction of Convolutional neural network: according to the characteristics of Folk art images, choose the appropriate Convolutional neural network structure. For example, VGG, ResNet, Inception, etc., and adjust and optimize the network based on actual situations. Division of training set and verification set: divide the collected Folk art image set into training set and verification set for training and verification of network model. Selection of optimizer and Loss function: select suitable optimizer and Loss function to optimize network model parameters and adjust network structure to improve detection accuracy and robustness. Model evaluation and parameter tuning: Evaluate the performance of the classifier model by calculating indicators such as detection accuracy, accuracy, and recall, and perform parameter tuning and model improvement. Implementation of real-time detection system: the trained classifier model is applied to the real-time detection system to achieve automatic detection and recognition of Folk art images.

In the process of using traditional CNN to classify images, with the layer-by-layer convolution and pooling of image information, the underlying feature information will be lost, and at the same time, due to the single convolution kernel, the image feature information is not fully extracted. In order to effectively use the bottom details of the image and extract the multi-scale feature information of the image, this paper presents an improved CNN based on SAA. The CNN model of art image recognition is shown in Figure 2.

The underlying logic of the CNN art image recognition model is mainly based on convolutional operations and local perception. Convolutional operations can recognize features such as edges and textures in images, while local perception can reduce the number of weights and improve the efficiency of the model. CNN usually includes convolution layer, pooling layer, Activation function, full connection layer and other components. Among them, the convolution layer extracts features from the input image, the pooling layer down samples the features to reduce the output size, the Activation function adds nonlinear features, and the full connection layer combines the features output by the convolution layer and the pooling layer, and outputs the classification results. When training CNN models, optimization algorithms such as gradient descent are usually used to minimize the Loss function, and Backpropagation is used to update model parameters.

The underlying features of the image represent the most primitive and basic essential visual characteristics of the image, such as color distribution, surface texture, target shape, etc. Usually, the original pixels of the image are analyzed and modeled to complete feature extraction. For new data and tasks, to design an effective feature often requires information in the corresponding field, and the original feature may not be applicable. Therefore, feature learning has been paid more attention by researchers, especially the rise of deep learning has proved the availability of feature learning. The structure of CNN is to simulate the real biological neural network.

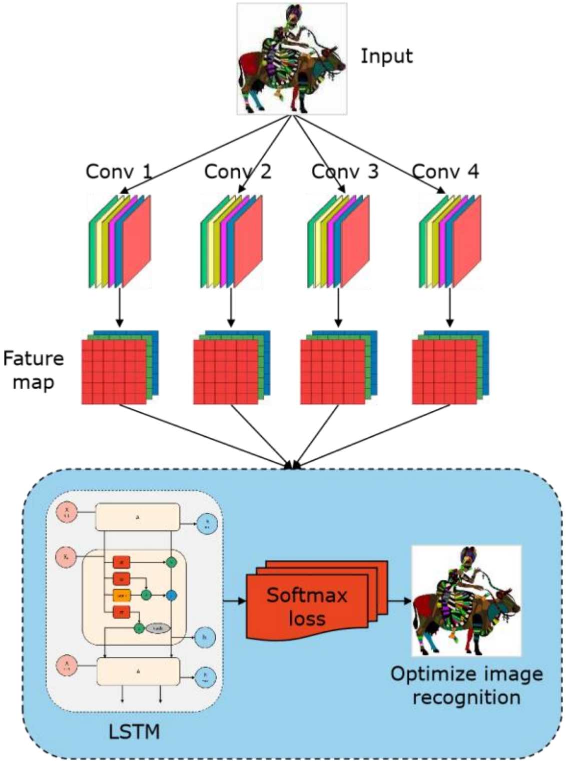

Figure 2: CNN model of art image recognition.

Data preprocessing directly affects the feature extraction of CNN. At present, the most common preprocessing method is to remove redundancy from data sets. On this basis, the average value of training samples is subtracted from the average value of samples, and the new samples obtained are the inputs of CNN. In the calculation process, if the input signal is $x \in R^{n \times m}$, the size of the convolution kernel is $w \in R^{s \times k}$. The resulting output signal:

$$y = x * w \in R^{u \times v} \tag{1}$$

The size of the extracted features:

$$u = \left[\frac{n - s + 2 \cdot Zeropadding}{Stride} \right] + 1 \tag{2}$$

$$v = \left[\frac{m - k + 2 \cdot Zeropadding}{Stride} \right] + 1 \tag{3}$$

CNN is a kind of supervised learning. All training samples and test samples are labeled. Random gradient descent method is used for training. Through network model training, the same kind of image samples are distributed in the same space, while other image samples are distributed in different spaces. The image feature expression is obtained by feature mapping from various artificially designed bottom feature descriptions, which contains further abstract information. The feature learning of images is an attempt to obtain the deep attributes shared by images through continuous optimization learning from a large number of data.

3.2 Local Feature Principal Direction Estimation and Convolution Operation Process

Local feature principal direction estimation and convolution operation are two important concepts in computer vision. Convolutional operation is a widely used operation in image processing and deep learning, while local feature principal direction estimation is a method used to describe local features in an image. In convolutional operations, the convolutional kernel is usually a small filter used to slide and extract features in the image. Each convolutional kernel outputs a numerical value representing the strength of the features detected by the convolutional kernel at that location. By using different convolutional kernels at different locations, various features in the image can be captured. Local feature principal direction estimation is a method of analyzing local image structure. It determines the main feature directions in a local image region by analyzing the feature directions in that region. This can help computer vision systems better understand the content in images and improve the accuracy of image processing and recognition. In summary, convolution operation and local feature principal direction estimation are important concepts in computer vision. Convolutional operations are used to extract features from images, while local feature principal direction estimation helps computer vision systems better understand the content of the image.

Content based image retrieval largely relies on image feature extraction techniques, namely image feature representation. In image retrieval, Scene graph can help us to describe the specific semantic information in images in a standardized way. To this end, we define two abstract concepts: Scene graph (used to describe the scene) and Scene graph specific association (each object in the Scene graph corresponds to the area in the image). These information can be obtained through computer vision technology, Natural language processing technology and other methods. Then, match the constructed Scene graph with the Scene graph stored in the database to find a Scene graph similar to the query image. Finally, the image corresponding to the Scene graph most similar to the query image is returned according to the similarity score. The image retrieval framework based on Scene graph can effectively describe colorful scenes and improve the accuracy of image retrieval. Since the task of Supervised learning is simpler and easier to train than Unsupervised learning, Supervised learning is more effective in descriptor generation and feature expression tasks, and can produce better results in feature matching. The convolution operation process of Folk art image performance feature extraction is shown in Figure 3.

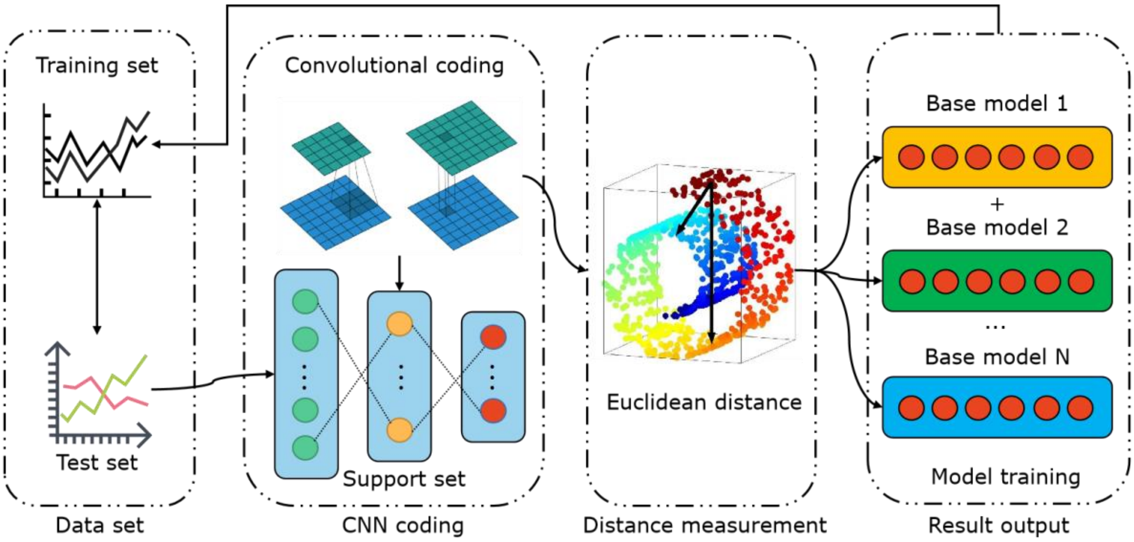

Figure 3: Convolution operation process of performance feature extraction of folk-art images.

As a powerful feature learner, the learning process of deep learning models is usually completed in an unsupervised and supervised manner. It can not only independently extract image features, but also relatively accurately learn the way humans understand images. The trained deep learning model can provide semantic descriptions similar to human visual systems for images, which is more conducive to the application of image detection and image retrieval. By incorporating deep CNN into the hash function, multi-level semantic similarity between multi label images is maintained, thereby avoiding the limitation of the semantic expression ability of manually designed features. This method can directly extract features from images during the frequency domain spatial decomposition process, making it suitable for general image systems that use texture content for indexing.

Users' aesthetics of folk-art products are no longer biased towards unchangeable design, but are more inclined to diversified and special-shaped product design. For a ceramic product image, $p(i)$ is the histogram probability of the image, i is the gray value of the image, $0 \le i \le L$. The histogram potential function is expressed as follows:

$$P_H(k) = \frac{1}{P_{\max}} \sum_{i=0}^{L} \frac{P(i)}{1+\alpha(i-k)^2} \tag{4}$$

$$P_{\max} = \max\left\{ \sum_{i=0}^{L} \frac{P(i)}{1+\alpha(i-k)^2} \right\} \tag{5}$$

Among them, α is a parameter. Let $I = \left[f(x,y) \right]_{m \times n}$ be an array of ceramic product images, where:

$$x \in \{0,1,2,3,\ldots,m-1\} \qquad y \in \{0,1,2,3,\ldots,n-1\} \tag{6}$$

$f(x,y) \in \{0,1,2,3,\ldots,G-1\}$ is the gray value of the pixel at the position (x,y) of the image array; G is the maximum gray value of the image I. The histogram function of image I is defined as:

$$h(i) = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \delta\left(f(x,y) - i \right) \qquad i \in \{0,1,2,3,\ldots,G-1\} \tag{7}$$

In the formula, the function $\delta(0) = 1$, $\delta(i \ne 0) = 0$. $h(i)$ represents the quantity of pixels whose gray value is i in the image I.

After the original painting data set is expanded by training samples, the classification accuracy of its CNN model has been significantly improved. The experimental phenomenon can be explained theoretically. On the one hand, the diversity of training samples is enhanced, which makes the training samples more abundant. When training the descriptor generation network, it is necessary to fix the network parameters of the main direction estimation part and only train the network parameters of the descriptor generation part. When both networks have been trained under their own conditions, that is, both networks have reached the best state under their own conditions, the main direction estimation network and descriptor generation network are jointly trained.

Assuming that the state vector sequence of the system is: $\{x_k, k \in N\}$, n_x is the dimension, and $x_k \in R^{n_x}$, which represents the state vector at time k, the state space model of the system is:

$$x_k = f_k\left(x_{k-1}, u_{k-1}, v_{k-1}\right) \tag{8}$$

Observation equation:

$$y_k = h_k(x_k, w_k) \tag{9}$$

Where y_k is the observed value of the system state vector, f_k and h_k are known nonlinear functions, the prior distribution of the initial state x_0 is $p(x_0)$, u_{k-1} is the input term, and v_k and ω_k are independent and identically distributed state noise and observation noise.

CNN has more hidden layer weight parameters, and its feature learning ability will be enhanced with the increase of training samples, thus improving CNN classification performance. The purpose of network joint training is to prevent the two networks from falling into their own local optima, thus failing to reach the global optima of the whole feature expression process. In the process of feature expression, the main direction estimation and descriptor generation network are combined into a whole network, and local image blocks are input and feature descriptors are output.

CNN can automatically learn efficient image feature representations through multi-layer convolution and pooling operations. Compared to manually designed features, CNN has a more powerful and flexible feature extraction capability. In art image classification tasks, CNN can capture key features in the image, such as edges, textures, shapes, etc., thereby improving classification accuracy. CNN can use different model structures and optimization algorithms for training and testing. In addition, methods such as data augmentation and data augmentation can be used to increase the amount of training data, thereby further improving the performance of the model. CNN can classify art images through multiple classifications. Compared with traditional classification algorithms, CNN has better robustness and can resist common image noise and interference, thereby improving classification accuracy. In the task of art image classification, it is necessary to collect a large number of art and non art images and annotate them in order to train an effective CNN model. In addition, the quality of annotated artistic images can also affect the performance of the model. CNN is sensitive to stylization of artistic images, and different types of artistic styles may have an impact on CNN's classification performance. The difference between the traditional western painting style and the modern Abstract art style may lead to the decline of the performance of CNN model. CNN performs well in art image detection and classification, with strong feature extraction capabilities and robustness, enabling accurate classification. However, the performance of CNN is influenced by the quality of annotated data and artistic style. In order to further improve the performance of CNN in art image classification, the following measures can be taken: collect more annotated art image data, improve data quality and diversity. Improve the structure of the CNN model to improve feature extraction ability and classification accuracy. Using Transfer learning and other methods, the pre trained CNN model is used to reduce training time and improve the performance of the model. Targeting different types of artistic styles to improve the generalization ability of CNN models.

4 RESULT ANALYSIS AND DISCUSSION

The classification of art images mainly studies the use of effective features extracted from art images to predict the categories of test samples. A small number of training samples means less effective information. At the same time, because of the subjective characteristics and artistic style of art images, its feature extraction and expression are more difficult than natural images, so the classification of small sample art images has always been a difficult point in art image detection. Convolutional neural network (CNN) can be used to classify artistic images. For art image classification tasks, CNN can effectively process image recognition tasks with translation, scaling and Rotational invariance, which are very important for art image recognition and analysis. In Convolutional neural network, the input image extracts features through the superposition of convolutional kernels. This process is called convolution operation. In art image classification tasks, the parameters of the convolutional kernel need to be adjusted through training to enable

CNN to extract image features optimally. The extracted features can be classified through a fully connected layer, thereby achieving the classification task of artistic images. It should be noted that Convolutional neural network needs a large number of training data and tags for training in the task of art image classification, which may lead to high costs. In addition, Convolutional neural network needs a lot of computing resources and time for model training and reasoning. The training of CNN is also facing difficulties. When training CNN models, a lot of labeled data is often needed to train the models to support the accuracy of the models. Even if there is a huge amount of data for training, it takes a lot of time to input a large number of samples corresponding to the very high demand for computing resources. Combined with the fact that there are small sample data sets in a specific field, it becomes a key point to choose a CNN that performs well in a small number of data sets. Due to the lack of general data set of folk art, the painting data set of this experiment comes from websites such as "Red Moving China". According to the painting categories, it is divided into four categories: ink painting, paper-cut works, prints and sketches, with 800 samples in each category, 700 training data and 100 test data. Firstly, the algorithm is trained by training samples, and the specific training loss of the algorithm is shown in Figure 4.

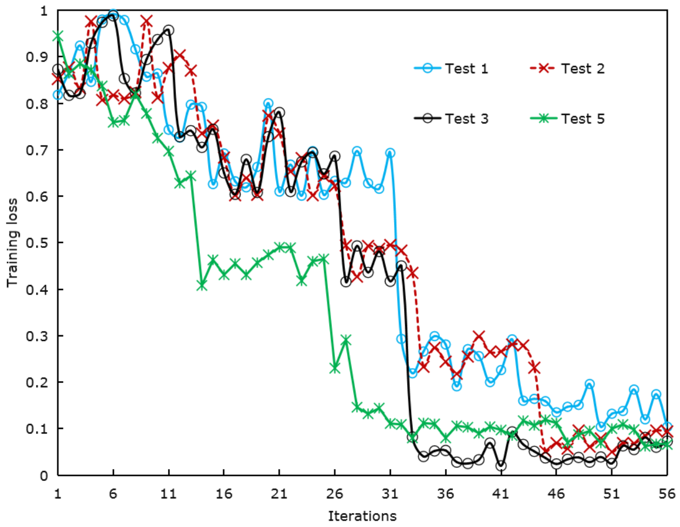

Figure 4: Loss of algorithm training.

It can be seen that the algorithm basically reaches convergence when the iteration is about 34 times. In the verification of DL algorithm, three evaluation indexes, namely, accuracy, mean absolute error and recommended time, are used for evaluation, in which mean absolute error represents the absolute mean of the difference between all individual observed values and actual values, and mainly judges the feature recognition error of the algorithm, which can directly reflect the feature recognition effect of the DL algorithm. Its expression is:

$$MAE = \frac{\sum_{C}^{i=1} \left| P_i^u - Q_i^u \right|}{c} \tag{10}$$

Where P_i^u stands for the predicted score made by user u on art work i, Q_i^u stands for the true score, and c stands for the total number of scores. The absolute error and the higher the recommendation accuracy. Accuracy represents the ratio of the number of successfully recommended results to the number of all recommended objects, and its expression is as follows:

$$\Pr ecision = \frac{R(u)}{T^{'}(u)} \times 100\% \tag{11}$$

Where $T^{'}(u)$ and $R(u)$ represent the number and feature quantity of all objects in the features calculated by the DL algorithm.

In the process of training the same batch of iterations, the same batch of samples will be randomly processed, and its output will change. In order to guide the direction of model parameter optimization, it is hoped that the output from the same input will be as same as possible, that is, the probability of belonging to a certain category is as close as possible for prediction classification. On the platform of Matlab, the efficiency of different classification models of artistic image performance characteristics is tested, and the recognition efficiency is evaluated by running time. The statistics of calculation time experimental results of different feature dimensionality reduction are shown in Table 1.

Image type	*Training sample*		*Test sample*	
	LSTM	*SAA-CNN*	*LSTM*	*SAA-CNN*
Chinese Brush Painting	7.38	6.37	8.25	5.91
Paper-cut works	8.33	6.68	6.88	4.76
Engraving	7.36	6.55	7.79	4.95
Sketch	6.85	5.43	7.43	5.69

Table 1: Dimension reduction time of classification model of art image performance characteristics.

To train an art image detection model using labeled painting samples, it is necessary to first collect a batch of image datasets containing art paintings and ordinary images. Then annotate these images, which can be divided into two categories: "art" and "non art". Professional annotation tools or programming annotation software can be used for annotation. Convolutional neural network (CNN) is used to extract the features of each image, and a fixed dimension feature vector is obtained. This feature vector is used as the feature representation of the image. The commonly used CNN models include VGG, ResNet, Inception, etc. The feature vectors of the pictures labeled as "art" are taken as positive samples, and the feature vectors of the pictures labeled as "non art" are taken as negative samples. Classifiers (such as Logistic regression, support vector machine, neural network, etc.) are used for training. During the training process, model parameters need to be adjusted to improve classification accuracy. Evaluate the trained model using a test set, calculate the accuracy, recall, F1 score, and other indicators of the model to evaluate its performance. Based on the evaluation results, optimize the model, such as adjusting model parameters, increasing the number of layers, and modifying the network structure. Apply the trained model to actual scenes to detect and recognize new artistic images. It should be noted that the training effectiveness of art image detection models is influenced by the quality and scale of the dataset, so it is necessary to collect high-quality and large-scale datasets as much as possible. In addition, the features of artistic images are subjective, so different people may extract different features, which may also affect the performance of the model.

For the problem of weak supervised learning, it is considered to train an art image detection model by using labeled painting samples, automatically label a large number of unlabeled art images through this classification model, get new labeled data, combine the original labeled data to form a new training set. Then, the classification model obtained by training is used to mine difficult samples, and the difficult samples in the training samples are screened out, which are added to the training samples in proportion and retrained to improve the classification accuracy of the classification model. If the dictionary constructed conforms to the characteristics of human visual system, the over-complete sparse representation model of the image can obtain a sparser image representation. The results of SAA-CNN and the comparison method for the classification model of artistic image performance characteristics are shown in Table 2.

Image type	*Training sample*		*Test sample*	
	LSTM	*SAA-CNN*	*LSTM*	*SAA-CNN*
Chinese Brush Painting	89.64%	95.35%	79.47%	93.46%
Paper-cut works	85.55%	96.28%	89.36%	94.17%
Engraving	87.42%	92.28%	88.29%	96.22%
Sketch	86.37%	94.36%	85.31%	93.33%

Table 2: Correct rate of classification model of art image performance characteristics.

Through continuous iterative optimization, the weight parameters of CNN classification model are constantly updated in the direction of image detection, thus continuously improving the accuracy of image detection by the classification model. Figure 5 shows the comparison results of the accuracy of art image feature recognition with different deep learning algorithms. Analysis of Figure 4 shows that the accuracy of the algorithm in this study is the highest, reaching about 95%, which is significantly higher than that of the traditional algorithm.

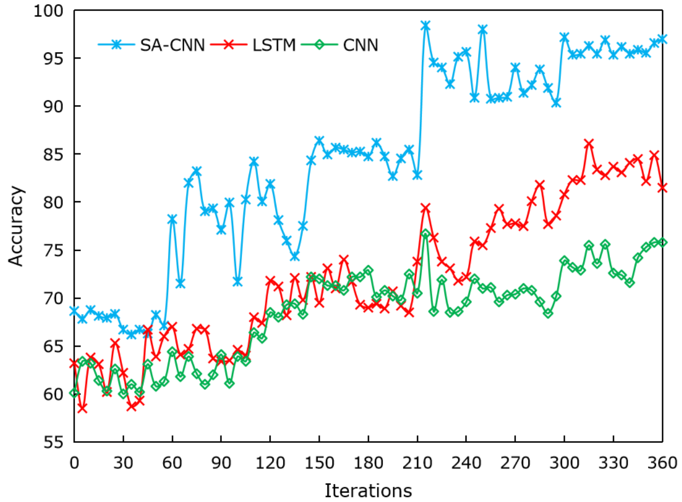

Figure 5: Comparison of recognition accuracy values of deep learning algorithm.

Because the gradient is varied in size and direction, and many different pixel distributions can produce the same gradient, if you want to use CNN to simulate the gradient successfully, you need more data for training. The art image analysis method for square absolute error in this study is superior to the vast majority of traditional algorithms, which can prove that the recommendation effect of the deep learning algorithm in this study is better.

Through Figure 6, it can also be found that the weight parameters of the natural image pre-training model with large data distribution correlation can be transferred to the art image detection model for learning, or the parameters of the art image pre-training model with identical training data can be directly transferred to the new art image detection model for further learning.

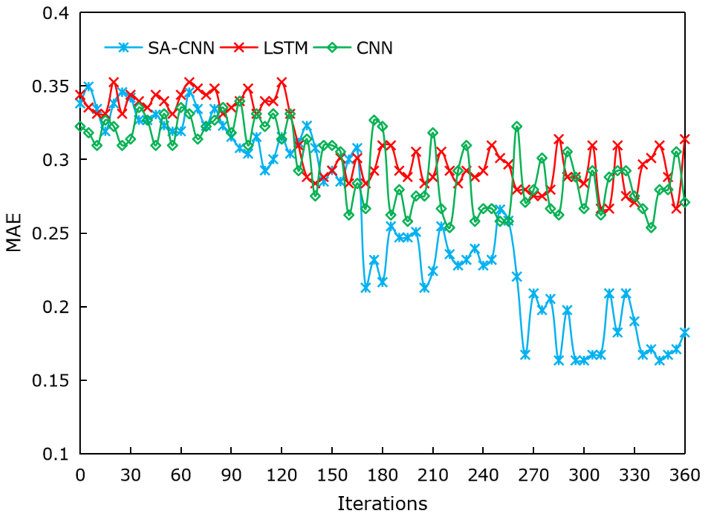

Figure 6: Comparison of average absolute error.

Images with multi-dimensional information characteristics of shape and texture combination will have different expression degrees under different classification algorithms. Different classification algorithms use different methods when selecting features. Some algorithms automatically select the most relevant features for classification, while others use all features for classification. Therefore, different algorithms may choose different features for classification, resulting in different levels of expression. Different classification algorithms have different model structures. Some algorithms use complex model structures to capture more features, while some algorithms use simple model structures to reduce overfitting. Therefore, different algorithms may capture different features, resulting in different levels of expression. Different classification algorithms use different datasets for training and testing. Some datasets may contain more noise or less useful information, while others may contain less noise or more useful information. Therefore, different datasets may lead to different levels of expression. Different classification algorithms have different parameter settings. Some algorithms may need more iterations or smaller Learning rate to achieve better performance, while others may not. Therefore, different parameter settings may lead to different levels of expression. When different classification algorithms process images with multidimensional information features, their expression level may be influenced by multiple factors. In practical applications, it is necessary to select appropriate classification algorithms and parameter settings based on specific problems and datasets to achieve the best classification results. Because of this, it is easy to make misclassification only by a certain feature, and the features obtained in this way are one-sided and cannot fully express the data features. The classifier receives the input feature vector of the object in the region of interest of the image, and finally outputs the category to which the target object belongs through the classification algorithm. The classifier calculates the similarity between the image to be classified and these categories through a set of classification algorithms, and the similarity value is a function value corresponding to the image feature vector, which can determine the classification. From the perspective of feature extraction, this paper makes full use of image texture features from the perspective of model training by using texture feature combination strategy. Through the comprehensive utilization of the texture information of the region of interest in the image, the strategy is optimized from two aspects: feature extraction and training model.

With the increase of task alienation, the possibility of feature transmission at higher level is greatly reduced. Therefore, the shared higher level may have the risk of reverse transfer, especially for the heterogeneous operation of the model. The image recognition time of the two methods is shown in Figure 7.

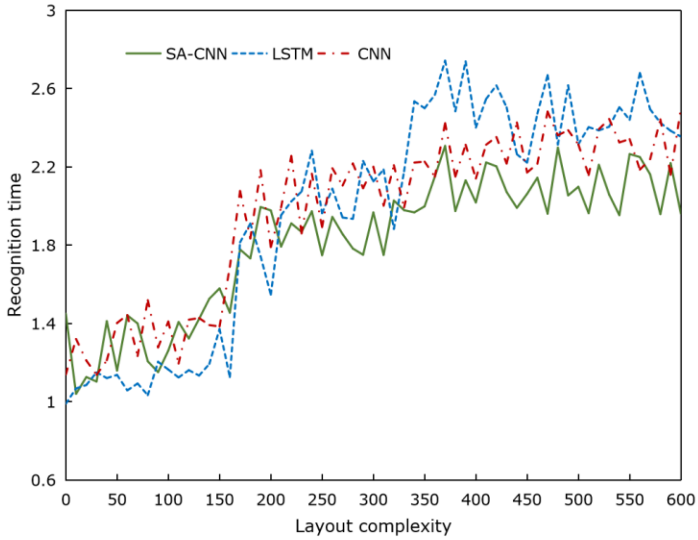

Figure 7: Time comparison of image recognition.

By analyzing Figure 7, it is found that the art image recognition method in this study takes less time in several experiments, which is less than the traditional method. Deep learning algorithm extracts features through multi-layer neural network, which can automatically learn efficient feature expression and avoid the problem of manual Feature engineering. This can capture complex feature information in the image, such as edges, textures, shapes, etc., thereby improving the accuracy of image recognition. Deep learning algorithms have flexible model structures and can design different network structures for different problems. For example, Convolutional neural network (CNN) is specially used for image recognition tasks and has good performance. By continuously improving the network structure, the accuracy of image recognition can be further improved. Through large-scale data training, deep learning methods can learn richer feature representations. At the same time, data enhancement and data increment can effectively alleviate the problem of Data deficient and further improve the accuracy of the algorithm. Deep learning algorithms typically use efficient optimization algorithms such as Random Gradient Descent (SGD) and Batch Gradient Descent (BGD), which can quickly update model parameters during the training process. At the same time, there are some improved optimization algorithms, such as Adaptive learning rate method and dropout technology, which can prevent over fitting and accelerate model training. In summary, deep learning algorithms have achieved significant results in the field of image recognition, mainly due to their powerful feature extraction ability, flexible model structure, large-scale data training, and efficient optimization algorithms.

To sum up, the deep learning algorithm in this study can achieve accurate image recognition in a short time, which is superior to the recommendation performance of traditional algorithms and can play an obvious advantage in CAD image feature recognition of folk-art images. CNN can extract rich and effective image features, so the feature expression method based on learning has better performance. The model parameters of natural image training and the model parameters of art image training have great sharing. The pre-trained model parameters of natural image training are migrated to the art image detection model, that is, the existing knowledge is migrated to new related fields, which is of great help to the art image data set lacking training samples, thus improving the classification accuracy of art images on the test set.

5 CONCLUSION

Folk art images record the growth of human traditional culture. The classification study of folk-art images is helpful to sort out painting resources, analyze the performance characteristics of folk art images, and better inherit China traditional culture. Art works are of great significance in the study of human history, science, art and culture, but it is difficult to obtain art works, which makes the research work of researchers more difficult. Traditional image detection mostly relies on manual extraction of features such as shape and color. Because art image detection needs more professional knowledge background, the process of manual extraction of features is complicated and complicated. This paper studies the application of CAD and deep learning technology in the feature analysis of folk-art images, and constructs an image feature extraction and classification model based on SAA-CNN. Compared with other traditional methods and deep learning methods, it is proved that the art image detection method based on CNN proposed in this paper has better classification performance for art images, which proves that this method is feasible and superior. For the traditional sparse representation model, the spatial flow pattern structure in the data is often not considered in the process of feature learning, which may affect the description of image features. Therefore, on the one hand, the follow-up work is to consider how to learn the image features directly from the original image content, rather than relying on the local features of the image.

Yulu Huo, https://orcid.org/0009-0006-9084-7468
Peng Zhao, https://orcid.org/0009-0009-2335-7559

REFERENCES

[1] Bappy, J.-H.; Simons, C.; Nataraj, L.; Manjunath, B.-S.; Roy, C.-A.-K.: Hybrid lstm and encoder–decoder architecture for detection of image forgeries, IEEE Transactions on Image Processing, 28(7), 2019, 3286-3300. https://doi.org/10.1109/TIP.2019.2895466

[2] Bayarri, V.; Sebastián, M.-A.; Ripoll, S.: Hyperspectral imaging techniques for the study, conservation and management of rock art, Applied Sciences, 9(23), 2019, 5011. https://doi.org/10.3390/app9235011

[3] Benzon, H.-H.; Chen, X.; Belcher, L.; Castro, O.; Branner, K.; Smit, J.: An Operational Image-Based Digital Twin for Large-Scale Structures, Applied Sciences, 12(7), 2022, 3216. https://doi.org/10.3390/app12073216

[4] Capotorto, S.; Lepore, M.; Varasano, A.: A Virtual Space Built on a Canvas Painting for an "Augmented" Experience to Catch the Artist's Message, ISPRS International Journal of Geo-Information, 10(10), 2021, 641. https://doi.org/10.3390/ijgi10100641

[5] Ding, D.; Yu, X.; Wang, Z.: The evolution of the living environment in suzhou in the ming and qing dynasties based on historical paintings, Journal on Computing and Cultural Heritage (JOCCH), 14(2), 2021, 1-14. https://doi.org/10.1145/3430700

[6] Gu, Y.; Wang, Y.; Li, Y.: A survey on deep learning-driven remote sensing image scene understanding: Scene classification, scene retrieval and scene-guided object detection, Applied Sciences, 9(10), 2019, 2110. https://doi.org/10.3390/app9102110

[7] Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J.: Graph convolutional networks for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing, 59(7), 2020, 5966-5978. https://doi.org/10.1109/TGRS.2020.3015157

[8] Hong, S.; Shen, J.; Lü, G.; Liu, X.; Mao, Y.; Sun, N.; Tang, L.: Aesthetic style transferring method based on deep neural network between Chinese landscape painting and classical private garden's virtual scenario, International Journal of Digital Earth, 16(1), 2023, 1491-1509. https://doi.org/10.1080/17538947.2023.2202422

[9] Jaspe, V.-A.; Ahsan, M.; Pintus, R.; Giachetti, A.; Marton, F.; Gobbetti, E.: Web-based exploration of annotated multi-layered relightable image models, Journal on Computing and Cultural Heritage (JOCCH), 14(2), 2021, 1-29. https://doi.org/10.1145/3430846

[10] McCartney, N.; Tynan, J.: Fashioning contemporary art: a new interdisciplinary aesthetics in art-design collaborations, Journal of Visual Art Practice, 20(1-2), 2021, 143-162. https://doi.org/10.1080/14702029.2021.1940454
[11] Murthy, C.-B.; Hashmi, M.-F.; Bokde, N.-D.; Geem, Z.-W.: Investigations of object detection in images/videos using various deep learning techniques and embedded platforms—A comprehensive review, Applied sciences, 10(9), 2020, 3280. https://doi.org/10.3390/app10093280
[12] Nordin, H.; Razak, B.-A.; Mokhtar, N.; Jamaludin, M.-F.: Feature Extraction of Mold Defects on Fine Arts Painting using Derivative Oriented Thresholding, Journal of Robotics, Networking and Artificial Life, 9(2), 2022, 192-201. https://doi.org/10.57417/jrnal.9.2_192
[13] Santoso, M.-H.; Larasati, D.-A.; Muhathir, M.: Wayang Image Classification Using MLP Method and GLCM Feature Extraction, Journal of Computer Science, Information Technology and Telecommunication Engineering, 1(2), 2020, 111-119. https://doi.org/10.30596/jcositte.v1i2.5131
[14] Xu, B.; Jiang, J.: Exploitation for multimedia asian information processing and artificial intelligence-based art design and teaching in colleges, ACM Transactions on Asian and Low-Resource Language Information Processing, 21(6), 2022, 1-18. https://doi.org/10.1145/3526219
[15] Xue, F.: Retracted Article: Innovative design of wall painting pattern based on microprocessor system and evolutionary computer technology, EURASIP Journal on Advances in Signal Processing, 2021(1), 2021, 100. https://doi.org/10.1186/s13634-021-00810-x
[16] Zhao, Y.; Samuel, R.-D.-J.; Manickam, A.: Research on the application of computer image processing technology in painting creation, Journal of Interconnection Networks, 22(Supp05), 2022, 2147020. https://doi.org/10.1142/S0219265921470204