International Journal of Machine Learning and Cybernetics (2020) 11:137–152
https://doi.org/10.1007/s13042-019-00963-0

ORIGINAL ARTICLE

Fine-art painting classification via two-channel dual path networks

Sheng-hua Zhong[1,2,3] · Xingsheng Huang[1,2] · Zhijiao Xiao[1,2]

Received: 1 July 2018 / Accepted: 1 May 2019 / Published online: 18 May 2019

Abstract

Fine-art painting expresses the state of mind and social culture of mankind. Automatic fine-art painting classification is an important task to assist the analysis of fine-art paintings. In this paper, we propose a novel two-channel dual path networks for the task of style, artist and genre classification on fine-art painting image. It includes the RGB and the brush stroke information channels. Besides the RGB information channel is used to represent the color information in fine-art painting images, the brush stroke information channel is used to extract brush stroke information from fine-art painting images. And the four-directional gray-level co-occurrence matrix is used in deep learning to detect the brush stroke information, which has never been considered in the task of fine-art painting classification. Experiments on two datasets demonstrate that the four-directional gray-level co-occurrence matrix is effective in feature representation of fine-art painting images. And the proposed model achieves best classification accuracy and good generalization performance when compared with other methods.

Keywords Image classification · Fine-art painting classification · Brush stroke · Gray-level co-occurrence · Dual path networks

1 Introduction

In the history of world civilization, fine-art painting plays a very important role [1]. A recognized definition of fine-art painting is a visual art considered to have been created primarily for aesthetic and intellectual purposes and judged for its beauty and meaningfulness [2]. Fine-art painting is portraying of an era by a painter, through which it can discover the social scene and cultural environment at that time [3]. Nowadays, smart mobile devices have penetrated into every detail of people's daily life, which leads to the rapid development of digital collection of fine-art painting. Hence, vast digital collections have been made available across the Internet and museums. With a large number of digital works collection, it is very important to automatically process and analyze the fine-art painting. Therefore, there are a lot of tasks based on it, such as forging detection [4], object retrieval [5, 6], archiving and retrieval of works of fine-art painting images [7, 8], etc. In these related tasks, automatically classify the style, artist, and genre can help people improve their understanding of fine-art painting and cultivate their aesthetic abilities. In this paper, we try to study the task of style, artist and genre classification on fine-art painting image.

For the classification task of images, traditional machine learning methods have been used. The features used included GIST, SIFT and so on [9]. The classifiers include SVM, Improved Fisher Encodings, etc [10]. Since 2012, the method of deep learning has been also widely used. AlexNet proposed by Krizhevsky significantly improved the effect of the CNN-based model on natural image classification, there has been a significant shift away from shallow image descriptors towards deep features [11]. Evidences from previous work show that deep learning methods success is relied on the availability of large-scale datasets with labels

✉ Zhijiao Xiao
cindyxzj@szu.edu.cn

Sheng-hua Zhong
csshzhong@szu.edu.cn

Xingsheng Huang
huangxingsheng2016@email.szu.edu.cn

1 College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518000, People's Republic of China

2 The National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University, Shenzhen 518000, People's Republic of China

3 Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University, Shenzhen 518060, People's Republic of China

[12–17]. For example, for the classification of ImageNet, AlexNet can achieve good performance [11]. One important reason is that the number of ImageNet and the number of parameters of AlexNet is consistent. According to the rule of 10, the amount of training data should be larger than ten times of the number of learning parameters in the model.

As we known, driven by the increases of depth, the notorious problem of vanishing/exploding gradients could hamper convergence of the deep networks. He et al. partially solved this problem by introducing a deep residual learning framework [18]. Huang et al. [19] proposed a different network architecture that achieved comparable accuracy with deep ResNet [18], named dense convolutional networks (DenseNet). One big advantage of DenseNet is their improved flow of information and gradients throughout the network, which makes them easy to train. Moreover, dense connections have a regularizing effect, which could reduce the overfitting problem on tasks with smaller training set sizes. The deep residual networks implicitly reuse the features through the residual path, while densely connected networks keep exploring new features through the densely connected path. Chen et al. proposed a novel dual path architecture, called the dual path networks (DPN) [20]. This new architecture inherits both the advantages of residual and densely connected paths, enabling effective feature re-usage and re-exploitation.

However, for the classification of fine-art paintings, direct use of deep learning based methods is not as good as expected. One of the main reasons is that the number of samples for fine-art painting classification is limited. For example, Painting-91, which is a currently public large-scale fine-art painting dataset [9], only has 4266 images. Therefore, considering the very limited training data, deep learning based methods are difficult to directly extract features and achieve good performance. Considering that the samples of the currently published dataset of fine-art painting images are very limited, a CNN model could be firstly pre-trained on a large-scale dataset such as ImageNet, and then it is fine-tuned with the target dataset. Thus, the fine-tuning can, in the case of the limited sample size of fine-art painting datasets, construct an effective learning model based on the pre-trained CNN. Thus, in this paper, our proposed model also uses ImageNet dataset to pre-train our model.

In the task of fine-art painting classification, although some researchers attempted to use some existing deep learning model or construct some new deep learning models, limited existing work took into account the essential characteristics of fine-art paintings, especially brush stroke. Brush stroke is an important and inherent characteristic of the painting [21]. Although this important character has been considered in the classification of fine-art painting, it is not well represented. Only two existing work tried to use the computational model to represent them. Johnson et al. used wavelets and hidden Markov models to trace the individual brush stroke information for artist identification [22]. Li et al. proposed a novel extraction method by exploiting an integration of edge detection. The textured patterns based on these extracted edges were utilized to distinguish van Gogh from his peers [23]. Although these two existing work achieved good results in artist identification, the size of the dataset is very limited and their models were constructed only for two-class classification problem (van Gogh vs. Non-van Gogh). In this paper, we attempt to use the gray-level co-occurrence matrix (GLCM) to represent brush stroke. And the four-directional gray-level co-occurrence matrix is used in deep learning to represent and detect the brush stroke information, which has never been considered in the task of fine-art painting classification.

In this paper, we propose a novel two-channel dual path networks for the classification task of fine-art paintings. This model is consisted of two channels, in addition to the RGB channel generally be used, we create the brush stroke information channel, which is used to extract the brush stroke information from fine-art painting images. This model firstly pre-trains on ImageNet dataset, and then it is fine-tuned with the fine-art painting dataset.

According to the main remarks highlighted above, we propose a fine-art painting via two-channel dual path deep learning networks model combined with brush stroke information to classify fine-art painting images. In particular, the main contributions of our work are:

- The four-directional gray-level co-occurrence matrix in deep learning is proposed to represent and detect the brush stroke information.
- A novel two-channel deep learning model is proposed with the brush stroke information channel and the RGB channel. Each channel contains the dual path networks consisting of residual path and densely connected path.
- The proposed model has achieved better classification performance than some state-of-the-art fine-art painting image classification methods and other deep learning structures.

The rest of this paper is organized as follows. The second section briefly introduces the related work for fine-art painting classification. The third section introduces the architecture of our proposed model. The fourth section introduces the experimental setting and provides the experimental results. Finally, the conclusion and future work are drawn in section five.

2 Related work

With the advancement of the machine learning methods in image classification [24, 25], there exist some attempts to tackle the classification of fine-art painting related problems.

Hentschel et al. proposed the improved Fisher encoding method, which achieved better results with less training data [10]. Saleh et al. proposed the features called as Classemes on painting images [26]. And the classification based on Classemes showed superior performance for the fine-art painting image classification [27]. Falomir et al. compared K-nearest neighbors (KNN) and support vector machine (SVM) classification methods on fine-art painting image classification tasks, and experiments demonstrated that KNN classifier outperformed SVM classifier [28].

Since AlexNet achieved amazing performance in ILSVRC2012, deep learning methods began to dominate the task of image classification [11]. He et al. proposed deep residual networks (ResNet) to solve the problem of vanishing/exploding gradients [18]. This is because ResNet can solve the degradation problem well and get better accuracy at deeper depths. The degradation problem suggests that the solvers might have difficulties in approximating identity mappings by multiple nonlinear layers. Huang et al. proposed dense convolutional networks (DenseNet) [19]. For each layer in DenseNet, the feature-maps of all layers were used as inputs, and its own feature-maps were used as inputs into all subsequent layers. One big advantage of DenseNet is their improved flow of information and gradients throughout the network, which makes the network easy to train.

In the classification task of fine-art painting images, some researchers simply employ the deep learning model as a feature extractor to extract the features of the image. Saleh et al. studied the influence of the relationship between different features and different types of metrics on the classification task of fine-art painting images [27]. Using these learned metrics to transform raw visual features into another space has significantly improved the classification performance of fine-art painting images. Lecoutre et al. investigated the use of deep residual neural to solve the problem of detecting the artistic style of a painting [29]. Their experiments demonstrated that 50-layer residual network performed best than other models. Sabatelli et al. explored the behavior of different convolutional neural networks that have been originally pre-trained on a very different classification task and shown how their performances can be improved when these networks are fine tuned [30]. Their experiments also showed that 50-layer residual network performed better than other deep convolutional architectures, i.e. Inception V3 and VGG 19. Some other researchers proposed novel models according to reconstructing the structure of CNN to improve the performance of the classification task with fine-art painting images. Peng et al. proposed cross-layer CNN, which is formed by cascading a number of modified CNN [31, 32], showing their efficacy on artistic style, artist, and architectural style classification. Each modified CNN in the cross-layer is as same as AlexNet except that some convolution layers are removed. Tan et al. replaced the last layer of CNN with a SVM classifier instead of a softmax layer, and experiments demonstrated that the performance of the modified version of CNN is better than that of original CNN [33].

Some researchers believe that only the RGB intensity information on the fine-art painting images itself is insufficient to represent the effective features of the fine-art painting images. Sun et al. proposed a CNN architecture with two pathways extracting color features and texture features, respectively [34]. The RGB pathway is the standard CNN architecture and the texture pathway intermixes the RGB pathway by integrating the Gram matrices of intermediate features in the RGB pathway. Similar with Gram matrix, the gray-level co-occurrence matrix (GLCM) is also considered as important features to represent brush stroke information and further classify the art works. Daec et al. used GLCM as features to assess the authenticity of the art works [35]. Inspired by the above works, Agarwal et al. considered four-directional GLCM as the features to classify fine-art painting images [36]. The dataset in their experiment only includes 4800 images. Experiments showed that the vectorized version of GLCM cannot achieve as good performance as other local features, i.e. SIFT. Cetinic et al. calculated four statistics (correlation, contrast, energy and homogeneity) of GLCM as features, and their experiments proved that the four statistics are better than other intensity or color features for image representation [37]. Although these work utilized gray-level co-occurrence matrix as the input to extract texture features for the task of fine-art painting classification, these existing work are not based on deep learning architecture.

In the conference version of our work [38], we proposed to use one-directional gray-level co-occurrence matrix as input to represent and extract the brush stroke information in the fine-art painting images. We utilized ResNet as the learning framework of the model. Through experiments, we found that the classification accuracy of our proposed two-channel model is obviously better than the model with only one channel. Different from the conference version of our work, in this paper, we propose a two-channel dual path model, which is called as fine-art painting via two-channel dual path networks (FPTD). The RGB channel is responsible for RGB information representation, the brush stoke information channel is responsible for representing and extracting brush stroke information. In the brush stroke information channel, the four-directional gray-level co-occurrence matrix rather than one-directional gray-level co-occurrence matrix is used as the input. Each channel contains the dual path networks consisting of residual path and densely connected path.

3 Fine-art painting via two-channel dual path networks

3.1 Two-channel dual path networks architecture

In this section, we introduce the details about the proposed model: fine-art painting via two-channel dual path networks (FPTD). The structure of the proposed FPTD is shown in Fig. 1. In RGB channel, the original RGB image of each image is input into the dual path networks (DPN) [20]. In brush stroke information channel, the GLCM image is used to represent the brush stroke information and is input into the DPN. We extract a 2688-dimensional vector/feature from each channel, and 2688 is the number of feature maps obtained from the last fully connected layer. Then, they are combined as a 5376-dimensional feature. This feature is the input to the SVM classifier.

3.2 RGB channel

The RGB channel uses the original fine-art painting image as the input to learn the model and classify the fine-art painting images. In addition, we extract a 2688-dimensional vector from the last fully connected layer in dual path networks of the RGB channel. In this paper, we use two versions of DPN, including 98 layers and 131 layers, inheriting from the original DPN, its structure is shown in Table 1. To the setting of building blocks, the number of blocks stacked, and the down-sampling stages, we follow the settings from [20]. Different from the original DPN, the size of the modified DPN is 12 rather than 3 in the third dimension of the first convolutional layer.

In Table 1, G refers to the number of groups, and k refers to the channels increment for the densely connected path. For the DPNs, we use $(+k)$ to indicate the width increment of the densely connected path. The overall design of DPN inherits backbone architecture of the vanilla ResNet/ResNeXt, which makes it easy to implement and apply to other tasks.

3.3 Brush stroke information channel

The brush stroke is a fundamental part of fine-art paintings, and it can also be used to analyze or classify fine-art paintings. Because the brush stroke is also known as the texture information of painting, we use gray-level co-occurrence matrix (GLCM) to describe this kind of information in fine-art painting images and it is utilized as the input of the brush stroke information channel.

Q is an operator that defines the relative position of two pixels relative to each other and consider an image $\mathbf{I}$ of size $M \times N$ with L possible gray levels. In detail, Q can also be thought of as n-by-2 array integers that specifying the distance between the pixel of interest and its neighbor, where n is the number of between the pixel of interest and its different neighbors. Each row in the array is a two-element vector, [row_offset, col_offset], that specifies the relationship, or offset, of a pair of pixels. Here, row_offset is the number of rows between the pixel-of-interest and its neighbor, while col_offset is the number of columns between the pixel-of-interest and its neighbor. Because the offset is often

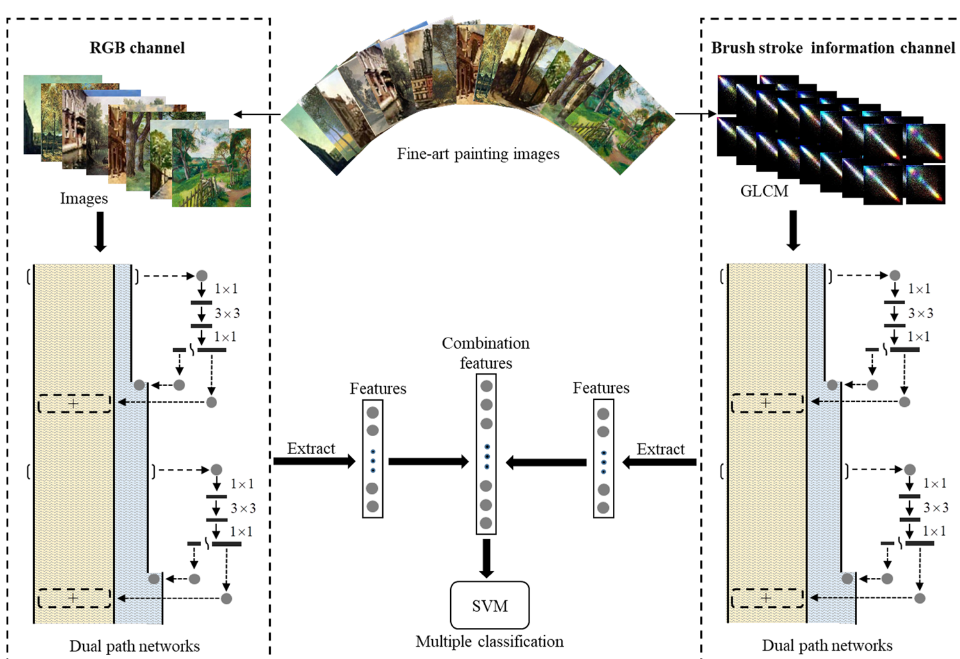


Fig. 1 Two-channel framework for fine-art paintings classification

Table 1 The architecture of DPN in our proposed model

Stage	Output size	98-layer DPN (40×4d)	131-layer DPN (40×4d)
Conv1	112×112	7×7×12×64, stride 2	
Conv2	56×56	3×3 max pool, stride 2	
		$\begin{bmatrix} 1\times 1,\ 160 \\ 3\times 3,\ 160, G=40 \\ 1\times 1,\ 256\ (+16) \end{bmatrix}\times 3$	$\begin{bmatrix} 1\times 1,\ 160 \\ 3\times 3,\ 160, G=40 \\ 1\times 1,\ 256\ (+16) \end{bmatrix}\times 4$
Conv3	28×28	$\begin{bmatrix} 1\times 1,\ 320 \\ 3\times 3,\ 320, G=40 \\ 1\times 1,\ 512\ (+32) \end{bmatrix}\times 6$	$\begin{bmatrix} 1\times 1,\ 320 \\ 3\times 3,\ 320, G=40 \\ 1\times 1,\ 512\ (+32) \end{bmatrix}\times 8$
Conv4	14×14	$\begin{bmatrix} 1\times 1,\ 640 \\ 3\times 3,\ 640, G=40 \\ 1\times 1,\ 1024\ (+32) \end{bmatrix}\times 20$	$\begin{bmatrix} 1\times 1,\ 640 \\ 3\times 3,\ 640, G=40 \\ 1\times 1,\ 1024\ (+32) \end{bmatrix}\times 28$
Conv5	7×7	$\begin{bmatrix} 1\times 1,\ 1280 \\ 3\times 3,\ 1280, G=40 \\ 1\times 1,\ 2048\ (+128) \end{bmatrix}\times 3$	$\begin{bmatrix} 1\times 1,\ 1280 \\ 3\times 3,\ 1280, G=40 \\ 1\times 1,\ 2048\ (+128) \end{bmatrix}\times 3$
	1×1	Average pool, 1000-dimensional fc, softmax	

The symbol $(+k)$ denotes the width increment on the densely connected path

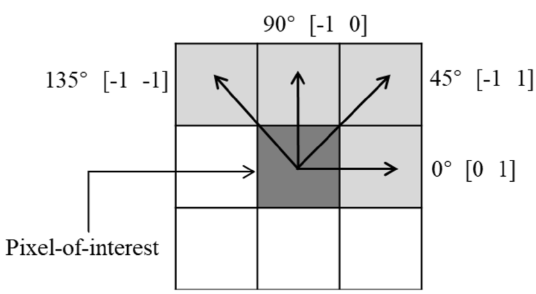


Fig. 2 The figure illustrates the array: offset = [0 1; − 1 1; − 1 0; − 1 − 1]

expressed as an angle, the following Fig. 2 shows the offset values that specify common angles, given the pixel distance being equal to 1. **G** is a matrix whose element g is the number of times the pixel pair with gray levels i and j appear at the position specified by Q in **I**, where $1 \leq i, j \leq L$.

When the angle of the direction is 0°, Q is defined as one pixel immediately to the right. Hence, **G** could be defined as Eq. 1:

$$\mathbf{G} = (g_{ij})_{L\times L}, \tag{1}$$

$$\begin{aligned} g_{ij} = |\{(x,y)|\mathbf{I}(x,y) = i,\ \mathbf{I}(x,y+1) = j, \\ 8\cdot m \leq x \leq 8\cdot m + 7,\ 8\cdot n \leq y \leq 8\cdot n + 7\}|, \end{aligned} \tag{2}$$

$$m = 0, 1, 2, \ldots, \left\lfloor \frac{M}{8} \right\rfloor + 1, \tag{3}$$

$$n = 0, 1, 2, \ldots, \left\lfloor \frac{N}{8} \right\rfloor + 1. \tag{4}$$

Similarly, it is also easy to calculate GLCMs in 45°, 90° and 135°. In our work, we obtain the gray-level co-occurrence matrix **G** for each color channel (R, G, and B). And then we combine them as a 3D matrix, which is referred as GLCM image.

Figure 3 shows four sample images with different styles and their corresponding gray-level co-occurrence matrix images, including 0°, 45°, 90°, and 135° directions. Figure 3a, c are sample images of "Impressionism". Figures 3e, g are sample images of "Post-Impressionism". Figure 3b, d are the corresponding four-directional GLCM image of Fig. 3a, c, respectively. Figure 3f, h are the corresponding four-directional GLCM image of Fig. 3e, g, respectively. We can find although these two different paintings are similar in vision, their four-directional GLCM images are different. The elements on the main diagonal in the GLCM are the number of occurrences of two pixels with the distance equal to 1 with respect to the grayscale combination. Since the color of the image is mainly continuously changed, the value of GLCM is concentrated near the main diagonal.

The brush stroke of fine-art painting mainly depends on the color, shape and thickness changes. As shown in Fig. 3a, the main color of this image is yellow. Thus, the main diagonal element of four-directional GLCM image shows yellow color. The sky part of Fig. 3a shows light blue color. But the intensity of the light blue is smaller than that of yellow. Therefore, the top left corner of the four-directional GLCM image, whose corresponding pixels are with smaller intensity, shows blue color. Moreover, compared with Fig. 3a, c, the brush strokes in Fig. 3e, g are much smoother. It makes the differences in RGB values of the adjacent pixels are smaller than that of Fig. 3a,

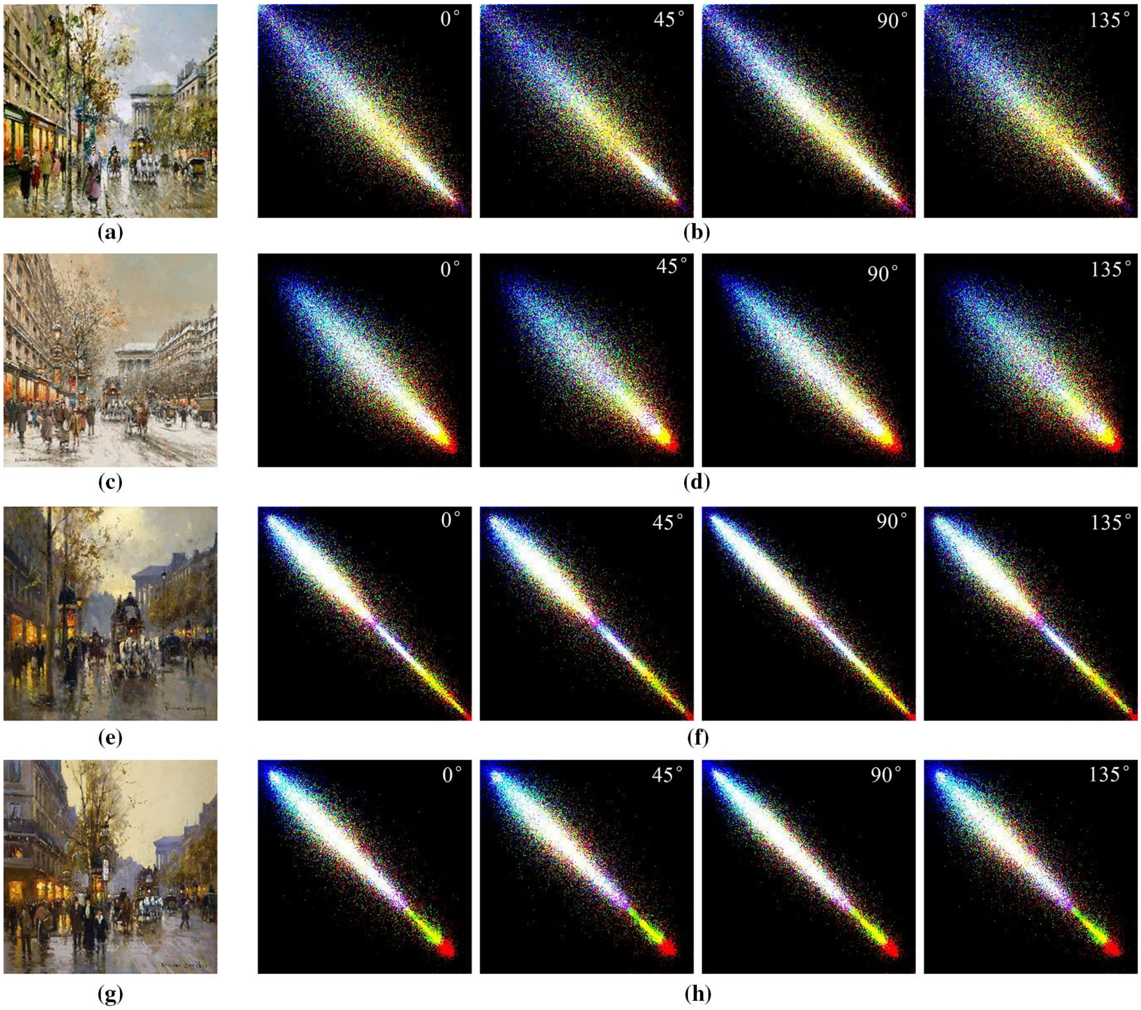


Fig. 3 **a** and **c** Sample images, which style is "Impressionism". **b** and **d** The extracted four-directional GLCM image of **a** and **c**, respectively. **e**) and **g** Sample images, which style is "Post-Impressionism". **f** and **h** The extracted four-directional GLCM image of **e** and **g**, respectively

c. From four-directional GLCM image, we can easily observe that the energy in Fig. 3e, g is more focused on the main diagonal.

This GLCM images are used as the input of brush stroke information channel and classify the fine-art painting images. In addition, we extract a 2688-dimensional vector from the last fully connected layer in dual path networks of the brush stroke information channel.

4 Experiment

In experiments, we first evaluate the proposed method on the WikiArt dataset in Sect. 4.1. Then, we test the generalization performance of the proposed method on the Gallerix dataset.

4.1 Experimental result of fine-art painting images classification

4.1.1 Experimental setting

In our experiment, we select the WikiArt dataset, which was frequently used in the classification task of art painting images, to validate the performance of the proposed method. The WikiArt dataset is the largest public digital fine-art painting image datasets, and it is available from WikiArt.org – Encyclopedia of fine-art painting images website. The fine-art painting images in this website as well as their annotations are contributed by a community of experts [10].

We collect three subsets of *style*, *genre*, and *artist* from the WikiArt dataset. The *style* subset includes a total of 30,825 fine-art painting images, including 25 styles. The *genre* subset has 28,760 fine-art painting images containing

10 genres. The *artist* subset has 9766 fine-art painting images, including 19 artists. The detail information of the specific subcategories in these three subsets is listed in Tables 6, 7, and 8. The resolution of the painting images in the three subsets varies greatly. Therefore, we resize all fine-art painting images to 256×256. In each dataset, 60% of fine-art painting images are used to train the compared model, and the remaining 40% are used for testing.

In order to illustrate the nuances of the fine-art painting images between different categories, Fig. 4 shows eight fine-art painting images belonging to different style categories in *style* subset. Although all of these images are forest, they belong to different style categories in fine-art painting images. Therefore, with respect to the difficulty of natural image classification, the classification for fine-art painting images is more difficult.

In our experiment, we propose to use GLCM to represent brush stroke information and compare it with Gram matrix [8]. In order to verify the validity of our proposed method, we compare other deep learning methods, such as AlexNet [11] and ResNet [18]. Following [20], we adopt standard data augmentation methods and train the networks using SGD with a mini-batch size of 32 for each GPU. For the deepest network, i.e. 131-layer DPN, the mini-batch size is set to 24. The learning rate starts from $\sqrt{0.1}$ for 131-layer DPN, and it starts from 0.4 for 98-layer, 50-layer and 14-layer DPN. It drops in a "steps" manner by a factor of 0.1. By following [11], the learning rate ε of AlexNet and ResNet for the training epoch p with respect to the current epoch i is set as,

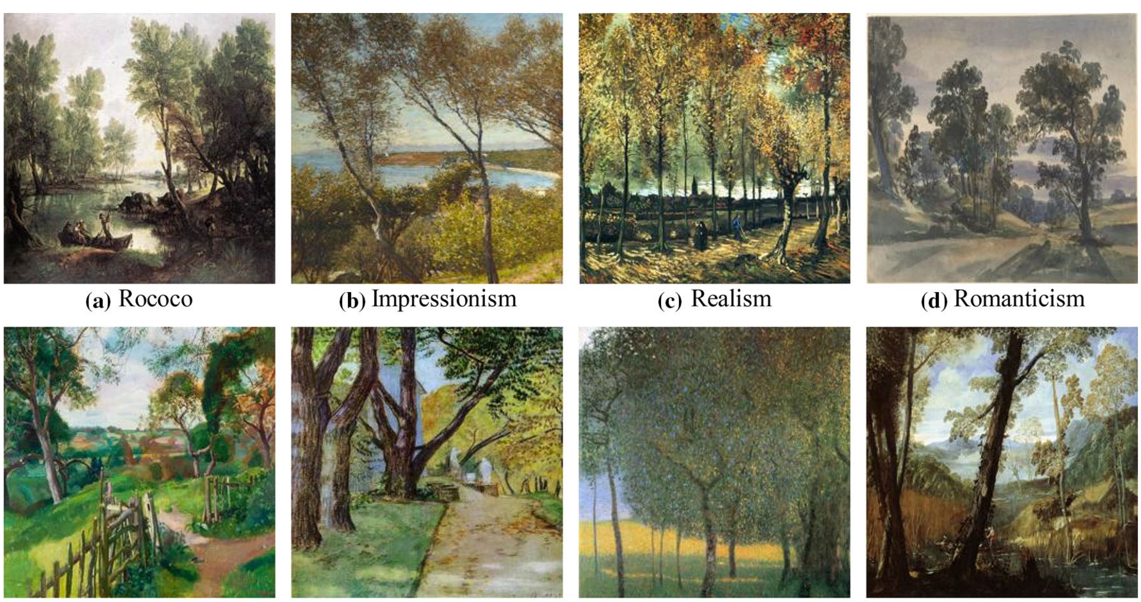
$$\varepsilon_i = 10^{-1-4\times\frac{i-1}{p-1}}, \tag{5}$$

where p is a positive integer to ensure that the model is convergent. In our experiments, when p is set to be 90, all the deep learning models are already converged.

We use LIBSVM Toolbox to implement SVM classifier and use the Gaussian kernel and the grid optimization to find the optimal value of C in the parameter space [2–10: 1000] with a step of 1 [27]. To overcome the limitation of the number of samples, our model firstly pre-trains on ImageNet dataset, and then it is fine-tuned with the fine-art painting dataset.

4.1.2 The effectiveness of the pre-training stage

The important findings of Tan et al. [33] show that the pre-training stage can effectively improve the performance of AlexNet in the classification task of fine-art painting images. Therefore, in the first experiment, we compare the performance of three deep learning models including AlexNet, ResNet, and DPN with or without pre-training stage. In the pre-training stage of this section, we pre-train the deep learning networks using images from ImageNet. Then, the fine-art painting images are used to fine-tune the pre-trained model. And here, only RGB information is included. We also compare the ResNet and DPN networks with four different depths, including14 layers, 50 layers, 98 layers and 131 layers. The detailed experimental results are shown in Table 2.



(a) Rococo **(b)** Impressionism **(c)** Realism **(d)** Romanticism

(e) Expressionism **(f)** Post-impressionism **(g)** Art Nouveau **(h)** Baroque

Fig. 4 Examples of images from the *style* subset which depict a similar scene but belong to different *style* categories

Table 2 The comparisons of the classification performance on *style*, *genre* and *artist* subsets using different deep learning models with or without pre-training stage

WikiArt's subsets	Networks	Without pre-training stage		With pre-training stage	
		Top-1 error Rate (%)	Top-5 error Rate (%)	Top-1 error Rate (%)	Top-5 error Rate (%)
Style	AlexNet [33]	69.23	31.76	56.71	19.12
	14-layer ResNet [18]	62.28	21.88	51.5	13.22
	50-layer ResNet [29]	67.16	27.08	49.91	11.96
	98-layer ResNet [18]	69.72	28.31	52.11	14.60
	131-layer ResNet [18]	71.88	30.15	53.51	15.08
	14-layer DPN [20]	**54.15**	**19.67**	47.83	10.97
	50-layer DPN [20]	55.39	20.42	46.42	10.23
	98-layer DPN [20]	56.89	21.97	**44.76**	**9.55**
	131-layer DPN [20]	60.46	24.38	44.96	9.68
Genre	AlexNet [33]	51.18	10.38	34.95	4.15
	14-layer ResNet [18]	48.65	7.78	32.91	3.43
	50-layer ResNet [29]	51.61	9.69	31.04	3.04
	98-layer ResNet [18]	53.47	10.57	31.40	3.09
	131-layer ResNet [18]	55.19	11.28	31.84	3.13
	14-layer DPN [20]	**41.48**	**5.74**	27.65	3.47
	50-layer DPN [20]	43.18	6.43	26.31	2.78
	98-layer DPN [20]	44.96	6.89	25.95	2.36
	131-layer DPN [20]	47.25	8.12	**25.32**	**1.82**
Artist	AlexNet [33]	53.74	19.28	27.34	5.6
	14-layer ResNet [18]	44.29	11.42	19.61	2.93
	50-layer ResNet [29]	57.82	19.33	18.13	2.75
	98-layer ResNet [18]	60.92	21.36	18.65	2.68
	131-layer ResNet [18]	65.34	23.51	19.92	3.26
	14-layer DPN [20]	**32.77**	**7.64**	16.35	1.39
	50-layer DPN [20]	35.18	8.19	15.97	1.17
	98-layer DPN [20]	36.58	9.35	15.62	1.12
	131-layer DPN [20]	40.36	11.23	**14.08**	**0.84**

By the comparison of the second row and third row of each subset in Table 2 with the first row, we can observe that the performance of ResNet is better than the performance of AlexNet. This is because ResNet can solve the problem of vanishing/exploding gradient better than AlexNet in the case of limited samples. Similarly, by comparing the performance of DPN (from the sixth row to the ninth row) and the performance of ResNet (from the second row to the fifth row), it is easily to observe that the performance of DPN is better than that of ResNet. This attributes to the fact that DPN combines ResNet and DenseNet together, which inherits the advantages of DenseNet in exploring new and useful features, while also retains the advantages of ResNet in solving the problem of vanishing/exploding gradients.

For each subset, it exists a significant improvement in the performance of any network with pre-training stage compared with that without pre-training stage. It shows that ResNet and DPN, like AlexNet, are also helpful for learning with pre-training stage. Furthermore, to those models without pre-training, we can find the shallow networks will achieve better performance. That is due to the samples of the *style*, *genre*, and *artist* subsets are limited. Hence, in the case of without pre-training, when the network structure is deep, the model will tend to be more susceptible to overfitting.

We have also provided the comparison on the computational cost of the proposed algorithm and other network architecture in Table 3. The comparisons are conducted using MNXnet on a single K80 graphic card and all training samples are cached into memory. As we known, the depth of deep learning networks is generally proportional to the computational cost. Therefore, ResNet with fewer layers is faster than DPN with more layers. But it is also observed that although the proposed model has two information channels, and each channel has two paths. The computational complexity is comparable to other methods.

Table 3 The model size and runtime for different networks

Method	Model size	Running time (single GPU)		
		Styles	Genre	Artist
AlexNet	179 MB	2 h	2 h	1 h
ResNet 14-layer	31 MB	8 h	7 h	3 h
ResNet 50-layer	90 MB	9 h	8 h	3.5 h
ResNet 98-layer	152 MB	17 h	16 h	6 h
ResNet 131-layer	189 MB	21 h	19 h	7 h
DPN 14-layer	24 MB	9 h	8 h	3 h
DPN 50-layer	42 MB	12 h	11 h	4 h
DPN 98-layer	226 MB	15 h	14 h	5 h
DPN 131-layer	294 MB	18 h	15 h	6 h

4.1.3 GLCM with deep learning structure vs. other brush stroke representation methods with different classifiers

In this section of the experiment, at first, we try to compare the classification performance of the deep learning models with different input information, including the proposed GLCM and the state-of-the-art brush stroke representation method: Gram matrix [8]. Gram matrix is used to represent the prototype texture to classify the fine-art painting images and achieve state-of-the-art results [39]. By following the setting described in Sect. 4.1.1, which has achieved better classification performance, we pre-train the deep learning models using the images from ImageNet. And then, we fine-tune the deep learning models on fine-art painting images with different representation methods, the proposed four-directional GLCM, the one-directional GLCM [38], and the Gram matrix.

The proposed four-directional GLCM represents each color channel of the original RGB image by four directions, including 0°, 45°, 90°, and 135°. To each direction, it calculates the gray-level co-occurrence matrix by 8×8 sub-block in the original RGB image. The one-directional GLCM only includes the information from 0° direction. For the original RGB image $\mathbf{I}$, $\mathbf{I} \cdot \mathbf{I}'$ is the Gram matrix. Being an orderless statistics of local stationary features, Gram matrix is also used as a texture descriptor to classify fine-art painting images [39]. The detailed experimental results are demonstrated in Fig. 5.

From Fig. 5, we can observe that the performances of the deep learning models with GLCM as the representation method are obviously better than those using Gram matrix as the representation method. And we can also find that the classification performance of 131-layer DPN is slightly better than that of the 50-layer ResNet. Moreover, compared with two versions of GLCM, the four-directional GLCM shows an improved performance over the one-directional GLCM. All these experimental results evidence that the GLCM is a better method to represent the brush stroke information.

Although we know GLCM is better than Gram matrix to represent the brush stoke information, we cannot presume that GLCM can play a better role with the deep learning structure. Therefore, in this experiment, we compare with other existing statistical methods [37] or machine learning methods [28] that use GLCM as their input.

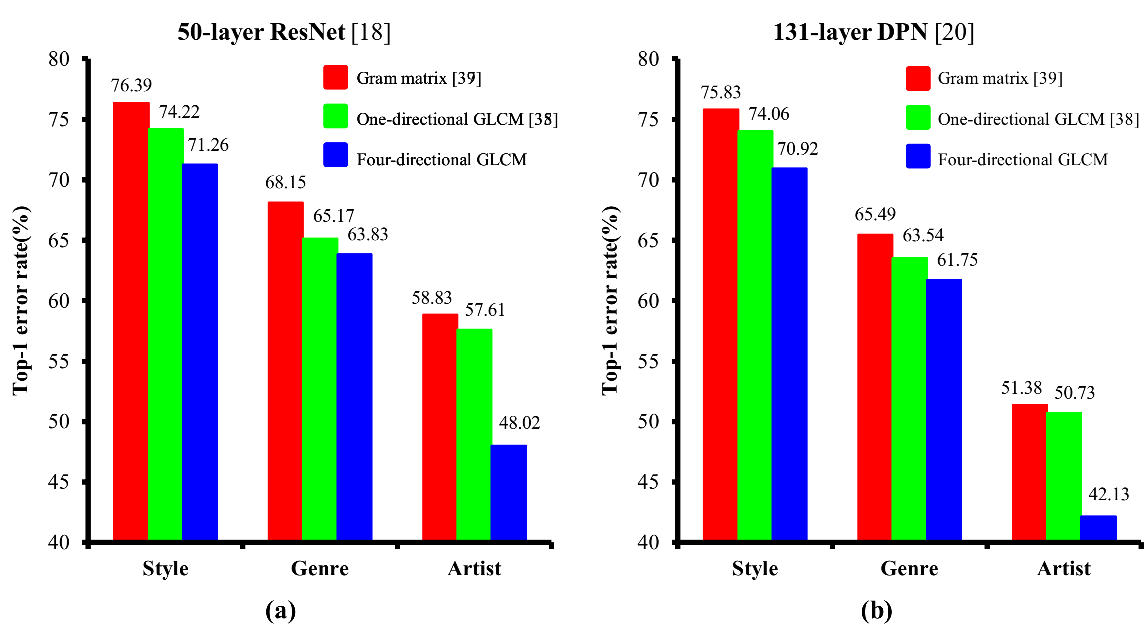


Fig. 5 The classification performance of fine-art painting images with different representation methods of the brush stroke information

Table 4 The Classification performance of fine-art painting images with the same input features but different classification methods

GLCM	Method	Top-1 error rate (%)		
		Styles	Genre	Artist
Four-directional	SMO [37]	**82.82**	**71.64**	**71.73**
	SVM [28]	98.40	96.00	97.90
	KNN [28]	98.40	96.00	97.90
One-directional	SMO [37]	91.08	80.28	84.31
	SVM [28]	98.40	96.00	97.90
	KNN [28]	98.40	96.00	97.90

In Table 4, the proposed four-directional GLCM and One-directional GLCM are identical to those of Fig. 5. The SMO method refers to the classification performance of sequential minimal optimization for support vector machine (SMO) based on four statistics (correlation, contrast, energy, homogeneity) of GLCM [37]. The SVM and KNN methods refer to the classification methods based on the SVM and KNN classifiers [28].

The experimental results of the first and fourth rows in Table 4 show that the classification performance of four-directional GLCM is better than that of one-directional classification. The performances via SMO are better than those via SVM and KNN. We suspect that is because expanding the high-dimensional GLCM into a vector would lose a certain degree of texture information, which ultimately causes SVM and KNN methods failure. More important, although all of them use GLCM as the input, we find the performances of these classification methods are worse than those with deep learning method.

4.1.4 The classification experiments based on two information channels: consistent pre-train information is important

In Sect. 4.1.3, we have verified that in the classification only with the brush stoke channel, four-directional GLCM as the input could achieve better performance. In this section, we integrate the brush stoke channel into two-channel model, to explore whether the performance on two-channel model is consistent with that on the brush stoke channel. In two-channel models, the original RGB images are used as the input of the RGB channel, while GLCM is used as the input of the brush stoke channel. Two versions of GLCM are included in the comparisons, including 50-layer ResNet and 131-layer DPN. The detailed experimental results are shown in Fig. 6.

From Fig. 6, we can find the 131-layer DPN shows lower error rate when it compares with 50-layer ResNet. This result is consistent with the existing results when the classification is on the brush stoke channel shown in Sect. 4.1.3. Moreover, we can also find the performances of the two-channel models are better than that of brush stoke channel. But when we compare the performance of the models with different representation methods, we obtain contradictory results from previous studies. Although the four-directional GLCM as the representation method achieves better performance in the classification on the brush stoke channel, it cannot has the similar effect in two-channel model. One-directional GLCM

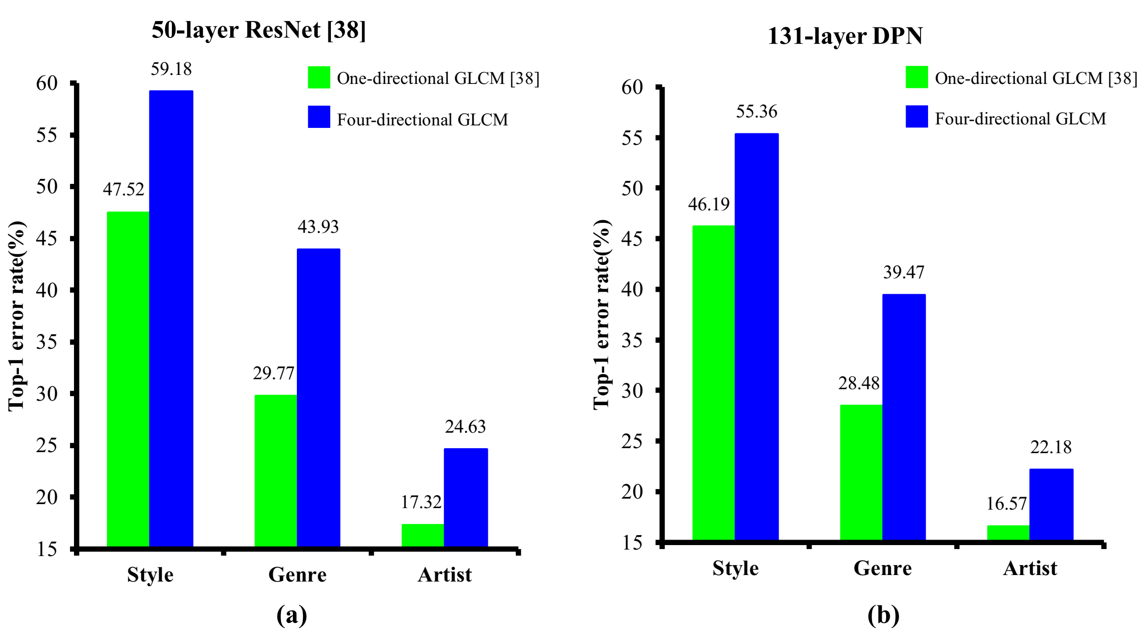


Fig. 6 The classification performance of the two channels of 50-layer ResNet and the 131-layer FPTD under different GLCM

obviously leads to better performance than four-directional GLCM.

We suspect it is probably due to the following two possible reasons. One possible reason is because the features extracted by the four-directional GLCM in the brush stroke channel are not an effective complementary to the features extracted by the RGB image in the RGB channel. Another possible reason is that the input of the brush stroke channel in the pre-training stage is inconsistent with the fine-tuned data. Specifically, the input of the brush stroke channel in the pre-training stage is the RGB images of ImageNet, while the fine-tuned data are different versions of gray-level co-occurrence matrix. As we know, four-directional GLCM information as the input will lead to more parameters of the deep learning models, and it will also result into high computational complexity. The inconsistent data content between pre-training and fine-tuning stages will make this problem even worse.

To rule out the influence of the inconsistent data content between the pre-training stage and the fine-tuning stage, we run another experiment and show its results in Table 5. Specifically, we transform the original RGB images of ImageNet to the gray-level co-occurrence matrixes as the input of the brush stroke channel in the pre-training stage. These results are reported with each kind of deep learning model in Table 5. To make the comparisons more clear, we also provide the existing results of the inconsistent setting in Table 5 when the input of pre-training stage is ImageNet. In Tables 6, 7, and 8, we provide the top-1 error rate (%) under different two-channel deep learning methods for each subcategory.

From these results, it is easily observed that the consistent pre-train information contributes the performance improvement for each model. Moreover, in consistent cases, four-directional GLCM is a better choice to represent brush stroke information when it compared with one-directional GLCM. In all compared model, the proposed FPTD achieves best performance.

4.2 The experimental results in generalization test

In this section, we try to validate the generalization ability of the proposed deep learning model on the Gallerix dataset, which is another famous collection of fine-art painting images.

4.2.1 The experimental setting

The fine-art painting images in the Gallerix dataset were downloaded from gallerix.asia. On this site, there are more than 160,000 paintings from world famous painters. This site is also thought to be one of the most complete collections of senior masters on the Internet. To keep consistent with the experimental setting of WikiArt, we also select the same subsets from the Gallerix dataset, i.e. *style*, *genre*, and *artist.* The *style* subset contains 20 classes with 20,411 images.

Table 5 The classification performance under different settings of brush stroke information channel

Experiment setting				The WikiArt of top-1 error rate (%)		
Channels	The setting of brush stroke information channel			Style	Genre	Artist
	Deep learning model	The input of ImageNet in pre-training stage	Fine-tuning dataset (WikiArt)			
Brush stroke information channel	50-layer ResNet [18]	Original images	One-directional GLCM images	74.22	65.17	57.61
		Original images	Four-directional GLCM images	71.26	63.53	46.98
		One-directional GLCM images	One-directional GLCM images	71.75	60.45	48.57
		Four-directional GLCM images	Four-directional GLCM images	66.2	53.74	37.84
	131-layer DPN [20]	Original images	One-directional GLCM images	74.06	63.54	50.73
		Original images	Four-directional GLCM images	70.92	61.75	42.13
		One-directional GLCM images	One-directional GLCM images	69.59	58.97	40.67
		Four-directional GLCM images	Four-directional GLCM images	**63.69**	**52.27**	**34.4**
Two-channels	50-layer ResNet [38]	Original images	One-directional GLCM images	47.52	29.77	17.32
		Original images	Four-directional GLCM images	59.18	43.93	24.63
		One-directional GLCM images	One-directional GLCM images	51.71	34.1	16.53
		Four-directional GLCM images	Four-directional GLCM images	51.26	33.31	16.09
	131-layer FPTD	Original images	One-directional GLCM images	46.19	28.48	16.57
		Original images	Four-directional GLCM images	55.36	39.47	22.18
		One-directional GLCM images	One-directional GLCM images	43.78	25.04	12.92
		Four-directional GLCM images	Four-directional GLCM images	**41.01**	**23.73**	**11.62**

Table 6 The subcategories in the WikiArt *style* subset, the classification performances under different two-channel deep learning methods for each subcategory, and the number of images for each corresponding in Gallerix dataset, '–' means there is no images in this corresponding subcategory

Subcategory	Top-1 error rate (%)		Image number in Gallerix dataset
	Two-channel deep learning methods		
	50-layer ResNet [38]	131-layer FPTD	
Abstract expressionism	59.23	49.04	957
Naive art (primitivism)	58.54	50.01	883
Neoclassicism	46.25	40.35	756
Rococo	28.25	24.90	829
Northern renaissance	37.60	33.23	694
Cubism	45.42	36.68	556
Minimalism	32.24	33.49	541
Pop Art	41.06	34.47	518
Art informel	68.84	53.39	376
Abstract art	57.61	45.11	340
Impressionism	50.92	45.14	3424
Color field painting	36.33	18.83	305
Ukiyo-e	11.99	11.25	–
Mannerism (late renaissance)	43.70	38.68	–
High renaissance	43.56	39.59	–
Early renaissance	35.37	19.33	–
Conceptual art	47.67	34.35	–
Realism	60.77	65.80	2056
Romanticism	53.05	57.94	1469
Expressionism	69.57	57.19	1106
Post-impressionism	58.42	47.20	1314
Surrealism	54.38	56.60	1137
Art Nouvzeau (modern)	45.64	49.12	1284
Baroque	51.22	36.72	891
Symbolism	52.13	48.84	975

Table 7 The subcategories in the WikiArt *genre* subset, the classification performances under different two-channel deep learning methods for each subcategory, and the number of images for each corresponding in Gallerix dataset, '–' means there is no images in this corresponding subcategory

Subcategory	Top-1 error rate (%)		
	Two-channel deep learning methods		Image number in Gallerix dataset
	50-layer ResNet [38]	131-layer FPTD	
Abstract	16.17	21.70	1769
Cityscape	26.96	22.55	1346
Genre painting	53.48	42.99	1243
Illustration	42.87	37.21	876
Landscape	28.61	22.56	2084
Nude painting Nu	26.35	17.47	–
Portrait	24.78	16.10	2815
Religious painting	38.96	29.15	634
Sketch and study	28.26	21.80	548
Still life	11.57	8.03	476

The *genre* subset contains nine classes with a total of 11,791 images. The *artist* subset contains 18 classes with a total of 6433 images. And the number of each class in these three subsets is varied, ranging from tens to thousands. The detail information of the specific subcategories in Gallerix dataset is shown in Tables 6, 7, and 8.

Figure 7 shows the some samples of the painting works in WikiArt dataset and Gallerix dataset. Figure 7a–h were created by Van Gogh at different stages, including four early works from WikiArt dataset and four late works from Gallerix dataset. All of them belong to category Post-Impressionist in the style subset. Although all these painting works belong to the category Post-Impressionist, they are different from each other on visual level. Figure 7i–l belong to the category Symbolism and are included in the WikiArt dataset. Figure 7m–p belong to the category Symbolism and are included in the Gallerix dataset. All of them belong to category Symbolism in the *style* subset. Hence, although both of the images from WikiArt and the Gallerix datasets are fine-art painting images, it is difficult to generalize the learnt model from WikiArt to Gallerix.

Table 8 The subcategories in the WikiArt *artist* subset, the classification performances under different two-channel deep learning methods for each subcategory, and the number of images for each corresponding in Gallerix dataset, '–' means there is no images in this corresponding subcategory

Subcategory	Top-1 error rate (%)		
	Two-channel deep learning methods		Image number in Gallerix dataset
	50-layer ResNet [38]	131-layer FPTD	
Albrecht Durer	11.13	8.10	329
Boris Kustodiev	25.06	20.64	131
Camille Pissarro	21.99	16.33	350
Childe Hassam	12.83	7.87	37
Claude Monet	28.70	21.36	720
Edgar Degas	15.83	11.91	103
Eugene Boudin	5.64	2.81	46
Ivan Aivazovsky	3.89	2.32	64
John Singer Sargent	9.38	4.13	138
Marc Chagall	22.81	16.36	498
Martiros Sarian	19.38	13.99	37
Nicholas Roerich	6.74	5.96	919
Pierre Auguste Renoir	11.59	6.28	786
Pyotr Konchalovsky	20.46	16.35	411
Raphael Kirchner	22.86	10.57	–
Rembrandt	7.25	6.45	166
Salvador Dali	23.24	18.40	151
Vincent Van Gogh	26.22	20.82	933
Pablo-Picasso	14.26	10.60	614

4.2.2 Generalization test with different model architectures

In the generalization experiment, we provide the generalization performance with different versions of models. The detailed experimental results are shown in Table 9.

By comparing the classification performance of deep learning models on the WikiArt, it is easily to find that the generalization performances of these models on the Gallerix are consistent with our previous observation. To the experiments on RGB channel, the performance of ResNet is better than the performance of AlexNet, and the performance of DPN is better than that of ResNet. In addition, the deeper networks will achieve better performance. In the two-channel models, the 131-layer FPTD we proposed has achieved best performance than others.

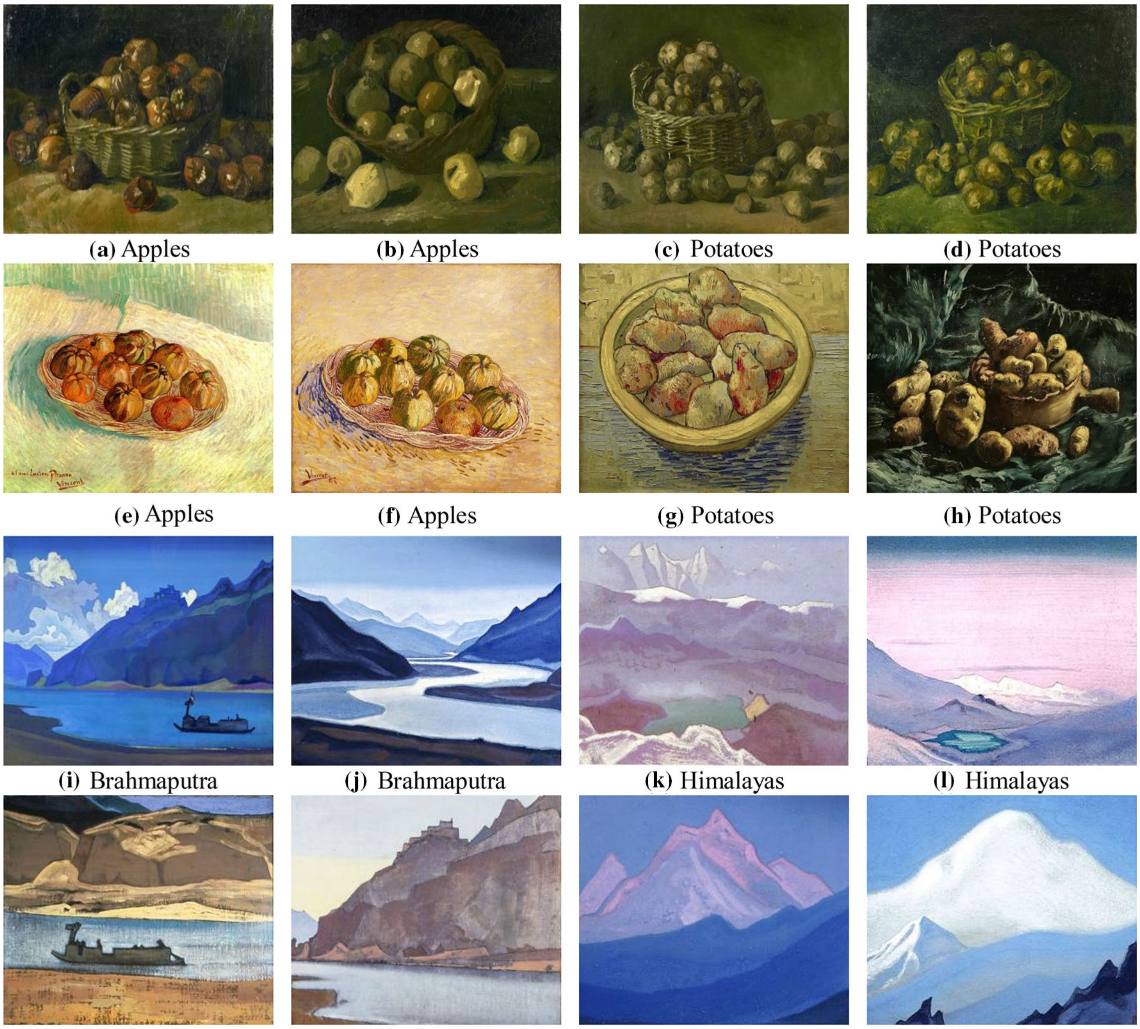
5 Conclusion and future works

Brush stroke is an inherent characteristic of the fine-art painting. In this paper, we propose a novel model for fine-art painting classification via two-channel dual path networks, including RBG and brush stroke information channels. In detail, we take the advantage of the ImageNet to pre-train the deep learning network. The gray-level co-occurrence matrix is used to represent the brush stroke information, as the input of the brush stroke information channel.

In order to validate the performance of our model, we run three experiments. In the first experiment, we find pre-training is helpful to learn an effective model. Among the performance of the three deep learning architectures, we also find that DPN has the best performance. Furthermore, to those models with pre-training, we can find the deeper networks of ResNet and DPN will achieve better performance. The second experimental results evidence that the GLCM is better than the Gram matrix to represent the brush stroke information. Moreover, the four-directional GLCM with deep learning structure shows an improved performance over other classification methods in the brush stroke information channel. In the third experiment, we first find some contradictory results obtained from two-channel models. That is, the one-directional GLCM as the input shows better performance over the four-directional GLCM. We speculate it is because the input in the pre-trained stage is not consistent with that in the fine-tuned stage for brush stoke information channel. And four-directional input information will lead to more parameters, and higher model complexity. The inconsistent data content between pre-training and fine-tuning stages will aggravate this problem. Therefore, in pre-train stage, we convert the original images to the corresponding GLCM as the input of the brush stoke information channel. After this transformation, the proposed two-channel model with four-directional GLCM achieves best performance than others.

In addition, we also provide the generalization test of the proposed method on another dataset of fine-art painting images. By comparing the classification performance of deep learning models on WikiArt, we can observe that the generalization performance of these models on Gallerix is similar.

In future, we will firstly try to integrate more characters of fine-art painting images into our model to improve the classification performance. Another meaningful future work is to propose a data augmentation method to address the issue of data shortage in fine-art painting images.



(**a**) Apples (**b**) Apples (**c**) Potatoes (**d**) Potatoes
(**e**) Apples (**f**) Apples (**g**) Potatoes (**h**) Potatoes
(**i**) Brahmaputra (**j**) Brahmaputra (**k**) Himalayas (**l**) Himalayas
(**m**) Brahmaputra (**n**) Brahmaputra (**o**) Himalayas (**p**) Himalayas

Fig. 7 **a**–**d** Belong to the Van Gogh's early works and are included in the WikiArt dataset. **e**–**h** Belong to Van Gogh's late works and are included in the Gallerix dataset. **a**–**h** Belong to the category Post-Impressionist in the *style* subset. **i**–**l** Belong to the category Symbolism and are included in the WikiArt dataset. **m**–**p** belong to the category Symbolism and are included in the Gallerix dataset. They all belong to the *style* subset

Table 9 Generalization performance of different models on Gallerix datasets

The model settings with fine-tuning		Top-1 error rate (%)					
		Training on WikiArt			Generalization on Gallerix		
Model architecture	Networks	Style	Genre	Artist	Style	Genre	Artist
Only RGB channel	AlexNet [33]	56.71	34.95	27.34	65.85	37.77	34.09
	14-layer ResNet [18]	51.5	32.91	19.61	60.13	31.11	25.59
	50-layer ResNet [18]	49.91	31.04	18.13	59.69	33.84	22.80
	98-layer DPN [20]	44.76	25.95	15.62	57.25	28.52	19.80
	131-layer DPN [20]	44.96	25.32	14.08	57.84	27.95	19.82
Two-channel deep learning methods	50-layer ResNet [38]	51.26	33.31	16.09	59.36	32.17	22.36
	131-layer FPTD	**41.01**	**23.73**	**13.26**	**49.47**	**26.14**	**17.73**

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. 61620106008), the Natural Science Foundation of Guangdong Province (Nos. 2016A030310053, 2016A030310039, 2017A030310521), the Science and Technology Innovation Commission of Shenzhen under Grant (No. JCYJ20160422151736824), the Shenzhen high-level overseas talents program, the National Engineering Laboratory for Big Data System Computing Technology, the Inlife-Handnet Open Fund, and the Postgraduate Innovation Development Fund Project of Shenzhen University (No. PIDFPZR2018001).

References

1. Štefko R, KrajĖák J (2013) An analytical view on fine arts marketing. Jerzy Kukuczka Academy of Physical Education Publishers Inc., Katowice
2. Alawa P (2017) Hegel and Gadamer on the contemporary understanding of art: an evaluation. Int J Philos Theol 5(2):58–65
3. Sultanova M, Mikhailova N, Amanzholova D (2016) Between the east and the west: Reflections on the contemporary art of Kazakhstan. Folklore 6:45–64
4. Polatkan G, Jafarpour S, Brasoveanu A, Hughes S, Daubechies I (2009) Detection of forgery in paintings using supervised learning. In: IEEE international conference on image processing (ICIP), pp 2921–2924
5. Crowley EJ, Zisserman A (2014) In search of art. In: European conference on computer vision (ECCV), pp 54–70
6. Crowley EJ, Zisserman A (2014) The state of the art: object retrieval in paintings using discriminative regions. In: British machine vision conference (BMVC), pp 1–12
7. Mensink T, Van Gemert J (2014) The Rijksmuseum challenge: Museumcentered visual recognition. In: Proceedings of international conference on multimedia retrieval (ICMR), pp 451–454
8. Gatys LA, Ecker AS, Bethge M (2015) A neural algorithm of artistic style. arXiv:1508.06576 **(preprint)**
9. Khan FS, Beigpour S, Van-De-Weijer J, Felsberg M (2014) Painting-91: a large scale database for computational painting categorization. Mach Vis Appl 25(6):1385–1397
10. Hentschel C, Wiradarma TP, Sack Harald (2016) Fine tuning CNNs with scarce training data-adapting ImageNet to art epoch classification. In: IEEE international conference on image processing (ICIP), pp 3693–3697
11. Krizhevsky A., Hinton G (2012) ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems (NIPS), pp 1097–1105
12. Zhong S, Liu Y, Hua K (2016) Field effect deep networks for image recognition with incomplete data. ACM Trans Multimed Comput Commun Appl (TOMM) 12(4):52
13. Zhong S, Liu Y, Li B, Long J (2015) Query-oriented unsupervised multi-document summarization via deep learning. Expert Syst Appl (ESWA) 42(21):8146–8155
14. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems (NIPS), pp 91–99
15. Banharnsakun A (2018) Towards improving the convolutional neural networks for deep learning using the distributed artificial bee colony method. Int J Mach Learn Cybern. https://doi.org/10.1007/s13042-018-0811-z
16. Pradhan R, Aygun RS, Maskey M, Ramachandran R, Daniel J (2018) Cecil.: tropical cyclone intensity estimation using a deep convolutional neural network. IEEE Trans Image Process 27(2):199–206
17. Xu X, Pan J, Zhang Y, Yang M (2018) Motion blur kernel estimation via deep learning. IEEE Trans Image Process 27(1):194–205
18. He K, Zhang X, Ren S, Jian S (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778
19. Huang G, Liu Z, Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: 30th IEEE conference on computer vision and pattern recognition (CVPR), pp 4700–4708
20. Chen Y, Li J, Xiao H, Jin X, Yan S, Feng JS (2017) Dual path networks. In: 31st conference on neural information processing systems (NIPS), pp 4467–4475
21. Li L (2011) The relationship between the brush strokes and the image, the color, the emotion. J Yangtze Univ 34(9):181–182
22. Johnson CR, Hendriks E, Berezhnoy I, Brevdo E, Hughes S, Daubechies I, Li J, Postma E, Wang JZ (2008) Image processing for artist identification—computerized analysis of Vincent Van Gogh's painting brushstrokes. IEEE Signal Process Mag Special Issue Visual Cult Heritage 25:37–48
23. Li J, Yao L, Hendriks E, Wang JZ (2012) Rhythmic brushstrokes distinguish Van Gogh from his contemporaries: Findings via automated brush stroke extraction. IEEE Trans Pattern Anal Mach Intell 34:1159–1176
24. Santosh KC, Roy PP (2018) Arrow detection in biomedical images using sequential classifier. Int J Mach Learn Cybern 9(6):993–1006
25. Huang R, Zhang G, Chen J (2018) Semi-supervised discriminant Isomap with application to visualization, image retrieval and classification. Int J Mach Learn Cybern 2018:1–10. https://doi.org/10.1007/s13042-018-0809-6
26. Torresani L, Szummer M, Fitzgibbon A (2010) Efficient object category recognition using classemes. In: European conference on computer vision (ECCV), pp 776–789
27. Saleh B, Elgammal A (2015) Large-scale classification of fine-art paintings: learning the right metric on the right feature. Int J Digital Art Hist 2:70–93
28. Falomir Z, Museros L, Sanz I, Gonzalez-Abril L (2018) Categorizing paintings in art styles based on qualitative color descriptors, quantitative global features and machine learning (QArt-Learn). Expert Syst Appl 97:83–94
29. Lecoutre A, Negrevergne B, Yger F (2017) Recognizing art style automatically in painting with deep learning. In: Asian conference on machine learning (ACML), pp 327–342
30. Sabatelli M, Kestemont M, Daelemans W, Geurts P (2018) Deep transfer learning for art classification problems. In: European conference on computer vision (ECCV), pp 1–16.
31. Peng K, Chen T (2015) Cross-layer features in convolutional neural networks for generic classification tasks. In: IEEE international conference on image processing (ICIP), pp 3057–3061
32. Peng K, Chen T (2015) A framework of extracting multi-scale features using multiple convolutional neural networks. In: International conference on multimedia and expo (ICME), pp 1–6
33. Tan W, Chan CS, Aguirre H, Tanaka, K (2016) Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification. In: IEEE international conference on image processing (ICIP), pp 3703–3707
34. Sun T, Wang Y, Yang J, Hu X (2017) Convolution neural networks with two pathways for image style recognition. IEEE Trans Image Process 26(9):4102–4113
35. Deac A. I, Lubbe J, Backer E (2006) Feature selection for paintings classification by optimal tree pruning. In: Multimedia content representation, classification and security, pp 354–361
36. Agarwal S, Karnick H, Pant N, Patel U (2015) Genre and style based painting classification. In: IEEE winter conference on applications of computer vision (WACV), pp 588–594

37. Cetinić E, Grgić S (2013) Automated painter recognition based on image feature extraction. In: Proceedings ELMAR-2013. Zadar, pp 19–22.
38. Huang X, Zhong S, Xiao Z (2017) Fine-art painting classification via two-channel deep residual network. In: Advances in multimedia information processing—Pacific-Rim conference on multimedia (PCM), pp 79–88
39. Ulyanov D, Lebedev V, Vedaldi A, Lempitsky V (2016) Texture networks: feed-forward synthesis of textures and stylized images. In: International conference on machine learning (ICML), pp 1349–1357