ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

Deep Learning in Computer Real-time Graphics and Image Using the Visual Effects of Non-photorealistic Rendering of Ink Painting

Li Wang[1], Jaewoong Kim[1,*]

[1]Graduate School of Advanced Imaging, Chung-Ang University, Seoul, 06974, South Korea
[*]Corresponding author: kjw@cau.ac.kr

Abstract: *This study is to explore the effect of deep learning algorithms applied to the non-photorealistic rendering (NPR) of ink painting. The powerful processing capabilities of NPR technology and convolutional neural networks (CNN) in computer vision are analyzed. Aiming at the features of ink texture images, a new sample-based ink texture synthesis method is proposed. The feature representation of ink texture is calculated using the CNN model, and then an ink texture with the same feature representation is generated on a random image. In addition, a synthesis method of ink painting picture is proposed based on the CNN. The results show that the ink texture generated by the ink texture synthesis algorithm is very good, and the Visual Geometry Group Network (VGG-Net) model shows the best generation effect. For images generated by high-level networks lack a lot of pixel information, but retain the main feature information of the image. When the learning rate is 10, the effect of the ink image synthesized based on the content image is slightly worse; and the image generated by the fusion of multiple styles confirms the feasibility of synthesizing ink images of different "sizes and styles". Therefore, this ink image synthesized on the basis of the CNN model shows a better effect, which extends the possibility of computer creation of richer ink painting works.*

Keywords: *non-photorealistic rendering technology; ink painting rendering; deep learning; convolutional neural network; computer graphic*

1. Introduction

Non-photorealistic rendering (NPR) is a research field that combines computer technology and painting art. It refers to using the computers to generate a technology that does not have a photo-realistic feeling, but has a hand-drawn style of graphics [1]. Compared with traditional realistic drawing and animation works, NPR focuses on simulating artistic works, expressing artistic characteristics and highlighting the details and features of local areas, which is more in line with people's aesthetic taste and has become a research hotspot in computer graphics. Chinese ink painting [2] is very different from the West in terms of drawing tools and techniques, and it implies the Chinese culture for thousands of years. Using a computer to synthesize an ink image with Chinese cultural characteristics [3] and integrating traditional art forms into computer technology-related applications can not only expand the scope of computer applications, but also promote the development of national culture.

Hauptfleisch et al. (2020) [4] proposed a new method to achieve real-time non-photorealistic rendering of a three-dimensional (3D) model, where the hand-drawn example specifies its appearance. The synthesis based on the guide patch is used to obtain high visual quality and time consistency. Multiple dimensions are considered to cover all degrees of freedom given by the available interaction space (e.g. camera rotation). Huang et al. (2019) [5] found that virtual brush modeling plays a vital role in virtual painting, and a powerful virtual brush model can truly reflect the features of real brushes and enhance the realism of virtual painting. Zheng et al. (2019) [6] proved that the artistic stylization of image is an important content of non-photorealistic rendering research, such non-realistic stylized research has a strong realistic romantic temperament and can well express the researcher's artistic pursuit. It is a good application example of the combination of computer science and art. In the image style conversion technology, the simulation of realistic art style mainly includes cartoon style, oil painting style, ink style, and pencil drawing style. At present, the simulation research on ink painting all over the world is relatively limited. Therefore, the research on Chinese ink painting method is not only conducive to the promotion of ink painting art, but beneficial for the development of computer digital media technology.

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

A model is constructed based on CNN and the actual synthesis effect is simulated in this study, the image effect of ink painting generated by the NPR technology in the computer is analyzed. The innovation of this study is to apply the CNN model in the deep learning algorithm to the NPR of ink painting, which provides a new idea for the visual effect of ink painting.

2. Research methods

2.1 NPR technology

NPR technology simulates the texture, color, line, and other style features of different works of art to generate an image with the style and aesthetic quality of the works of art [7]. Its goal is to formally specify a way to expand paintings, and then write computer programs that generate non-photorealistic paintings. It is closely related to disciplines such as computer science, digital media, and art design. NPR is distinguished from realistic rendering by its unique artistic expression methods, and shows the advantages of specific information transmission, assumptions about the image language, and a better understanding of the transmission mechanism [8]. Painters usually have intentions and emotions they want to convey when they create. They not only use different artistic methods, but also incorporate the painter's personal painting style. Therefore, it is hoped that the non-realistic works drawn by computers can show in a specific artistic way and can express certain emotions in the field of non-realistic graphics [9]. Therefore, NPR technology not only improves the efficiency of artistic production, but also narrows the distance between the public and art. It plays an important role in medicine, industry, and national cultural protection, and has attracted extensive attention from domestic and foreign researchers (Table 1).

Table 1: Comparison of NPR technology

Method	Image comparison	Stroke rendering	Image filtering	CNN	GAN	DCGAN
Type of style	Variety	Single	Single	Variety	Variety	Variety
Result	Not ideal	Not ideal	Not ideal	Excellent, it is often set as the standard of comparison	Good	Good
Drawing speed	Slow	Slow	Slow	Slow	Fast	Slightly high
Performance	Poor	Poor	Poor	Fast	Slightly fast	Slightly high

Note: “GAN” refers to “Generative Adversarial Networks”, and “DCGAN” represents “Deep Convolutional Generative Adversarial Networks”.

NPR is different from the previous photorealistic rendering methods. It is often used to express those painting forms that do not need to be realistic. It adopts artistic style to visually abstract the image, discards redundant details, focuses on certain features of the image, and simplifies the object. The shape highlights the information needed by the observer, making it easier for the observer to accept and understand [10].

2.2 Analysis on CNN and image texture

(1) CNN under deep learning

Deep learning originates from the study of neural networks and is a new research field in machine learning research. It forms an abstract high-level through the combination of low-level features to express the category of the attribute or the features of the attribute. Compared with traditional artificial neural networks (ANN), deep learning is not a simple shallow model, but a deep one [11]. The deep model can improve the accuracy of various machine learning algorithms, mainly because the model has the following two characteristics. First of all, natural hierarchization is its own hierarchical features. Shallow objects constitute middle-level objects, and middle-level objects constitute the whole. The various levels in the middle cannot be ignored or omitted. Secondly, bionics is a new subject proposed because biologists imitate the human nervous system. Biologists have discovered that the visual cortex of the human brain is hierarchical to the signal processing of the eye. They believe that the visual cortex first performs low-level abstract processing on an original eye input signal, gradually iterates to the high-level abstraction, and finally forms the brain's thinking.

Commonly used deep learning models include Auto Encoder, Sparse Coding, Deep Belief Networks (DBN), CNN, and Restricted Boltzmann Machine (RBM), among which CNN shows the best performance at dealing with computer vision [12]. CNN is a neural network that combines ANN and

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

deep learning methods, taking the error back propagation as the learning algorithm [13]. The propagation process is shown in Figure 1:

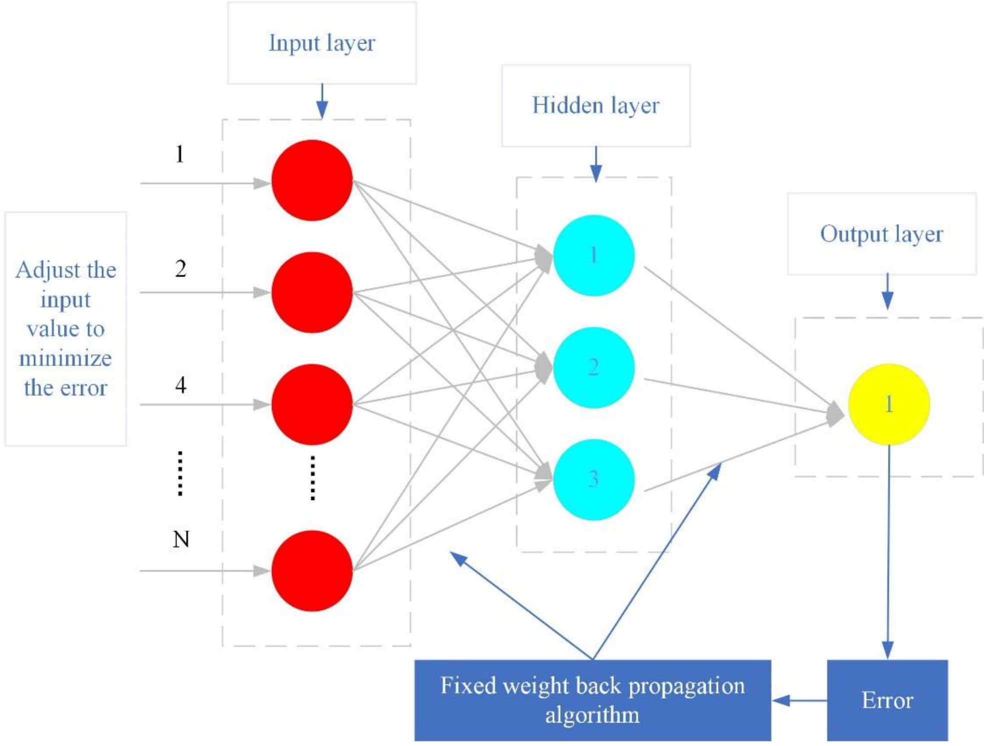

Figure 1: The error back propagation process of CNN.

A supervised training method is used for network convergence calculations. The error back propagation algorithm used in this study is not used to train the network, but to train the "source image". On the basis of the trained neural network model, the connection weights between the layers are fixed, the error signal is propagated forward layer by layer to the input layer, and then the input data is modified [14]. Compared with traditional ANN, CNN uses a convolutional layer that can extract features, and adds a down-sampling layer that reduces data dimensions but retains significant features, as shown in Figure 2:

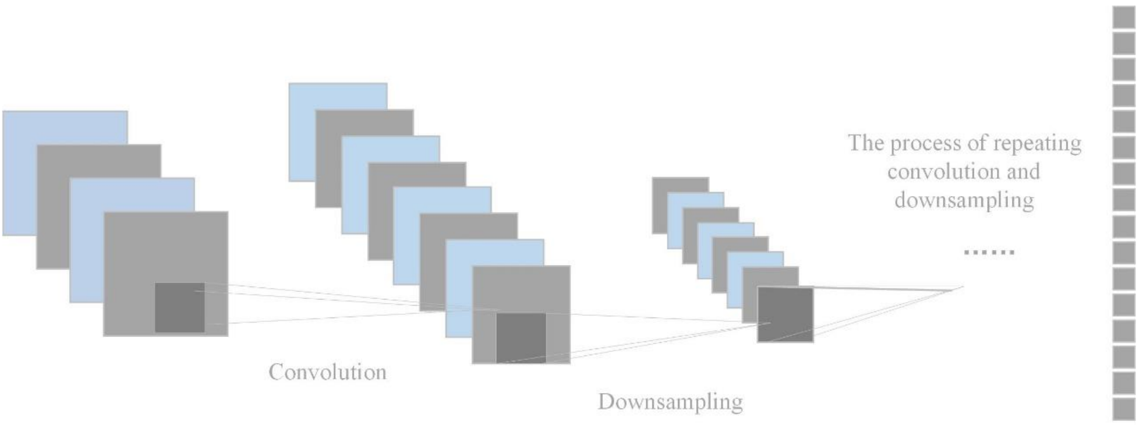

Figure 2: Basic structure of CNN.

In general, the first layer of CNN is a convolutional layer, which directly detects the features of the input image. As the network deepens, the convolutional layer is used to further detect image features, and the down-sampling layer is to eliminate redundant information. In addition, the output data of the output layer shows the same structure as the target tag information [15]. In addition to the structure of the traditional neural network, CNN shares the weights of neurons on the neuron plane, and extracts different features through different convolution kernels to ensure the reduction of computational complexity [16].

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

(2) Image texture synthesis

Texture synthesis is a method of constructing texture, which has a wide range of applications in all aspects of computer vision [17], and its goal is that the synthesized texture and source texture are visually unrecognizable. The use of texture synthesis technology can easily decorate the details of the virtual object model in computer graphics, without the need to separately model the surface of the object. Texture synthesis is also often used in the generation of large-scale scenes, NPR, and so on.

According to the different generation methods, the methods of synthesizing texture can be divided into three categories: texture mapping synthesis method, process-based synthesis method, and sample image-based synthesis method [18]. Texture mapping is to cover any plane figure or image on the surface of the object to form a real color pattern on the surface of the object, which is mainly used to deal with the surface aliasing. The process texture synthesis method is to synthesize the texture by analyzing the features of a certain texture and simulating its physical generation process. It is equivalent to writing a piece of code and using a computer to generate models such as smoke and ink [19]. The result of this method is very good, but every time a new texture is synthesized, and the parameters must be adjusted and tested repeatedly. The theoretical basis of texture synthesis based on sample images is to given a small piece of texture to generate a large piece of similar texture. Thus, it not only overcomes the shortcomings of the limited scope of the texture mapping method, but also avoids the cumbersome process of the process texture synthesis method.

2.3 Ink texture synthesis based on CNN

AlexNet [20] is a very typical CNN, which consists of 5 layers of convolutional layers, 2 layers of fully connected layers, and a label layer (with 1000 nodes, each node represents a category in ImageNet) [21]. The parameters used are 12 times less than Alex Net, but the accuracy rate is higher. In the field of object detection, the biggest contribution comes from the fusion of deep networks and classic computer vision algorithms. Visual Geometry Group Network (VGG-Net) [22] is a CNN developed from Alex-net. With the deepening of the network, VGG-Net tests CNN for large-scale image classification and positioning. VGG-Net uses a small convolution kernel and a small step size, and takes different ratios of image as input during training and testing. The trained CNN model is general for the extraction of image features, and the VGG-Net CNN has good image transformation learning capabilities. The purpose is to test the effect of CNN on large-scale image classification and localization as the depth of the network deepens.

The VGG-Net can be introduced as follow. The input image size is 224 × 224 × 3; the only preprocessing is to subtract the mean value from the image; the 1 × 1 convolution kernel can be understood as the linear transformation of the input channel; 3 ×3 size convolution kernel is more used; Max-Pooling is generally processed on a 2 × 2 pixel window with a step size of 2; and all layers use Rectificationnon-linearity (RELU) except the last fully connected classification layer; and do not add Local Response Normalization (LRN), because it does not improve the effect but will bring calculation and memory costs, and increase the calculation time. The following Table 2 shows the configuration of VGG-Net:

Table 2: Structural configuration of VGG-Net

<table>
<tr><td rowspan="6">ConvNet Configuration</td><td>A</td><td>11 weight layers</td><td rowspan="6">Input (224 * 224 RGB image)</td><td>Conv3-64</td><td rowspan="6">Max-pooling</td><td>Conv3-128</td></tr>
<tr><td>A-LRN</td><td>11 weight layers</td><td>Conv3-64
LRN</td><td>Conv3-128</td></tr>
<tr><td>B</td><td>13 weight layers</td><td>Conv3-64
Conv3-64</td><td>Conv3-128
Conv3-128</td></tr>
<tr><td>C</td><td>16 weight layers</td><td>Conv3-64
Conv3-64</td><td>Conv3-128
Conv3-128</td></tr>
<tr><td>D</td><td>16 weight layers</td><td>Conv3-64
Conv3-64</td><td>Conv3-128
Conv3-128</td></tr>
<tr><td>E</td><td>19 weight layers</td><td>Conv3-64
Conv3-64</td><td>Conv3-128
Conv3-128</td></tr>
</table>

In this study, features mapping space provided by the 16 convolutional layers and 5 pooling layers of VGG-Net are adopted without using any fully connected layer, so that the input image can be of any size. The first convolutional layer has the same size as the input image, and the ratio of the size of the features map between the subsequent network layers remains unchanged. Generally, a nonlinear filter bank is defined at each layer in the network, and its complexity increases as the depth of the network layer

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

increases. The trained network model can be more applied and researched under the framework of deep learning after simple modification and adjustment. For ordinary texture synthesis, replacing the maximum pooling of VGG-Net with average pooling processing can get a clearer result, but the maximum pooling processing is retained for synthetic ink texture.

Ink texture is a pattern that causes ink to diffuse into water, or a brush leaves a local regular pattern on the rice paper. Ink texture is a kind of natural texture, which is irregular locally but has certain regularity only on the whole. Early sample-based texture synthesis methods consider the idea of texture analysis too much, and use statistical methods and filters for processing [23]. The CNN is selected in this study to have a natural advantage in image features extraction, and a new ink texture synthesis method is proposed.

To extract texture from a given source ink painting texture image, it should extract features information of different sizes from this image uniformly firstly. Next, the image features response is calculated to get the spatial summary statistics of the description of a fixed source image. Finally, the gradient descent method is adopted on an image initialized with white noise, to generate an ink texture image with the same fixed description as the source texture ink image.

A high-performance DNN is adopted in this study instead of a linear filter and a set of carefully selected summary statistics to provide the features space, and there is only one spatial summary statistics. The relationship between the features responses of each layer of the network is shown in the Figure 3 below:

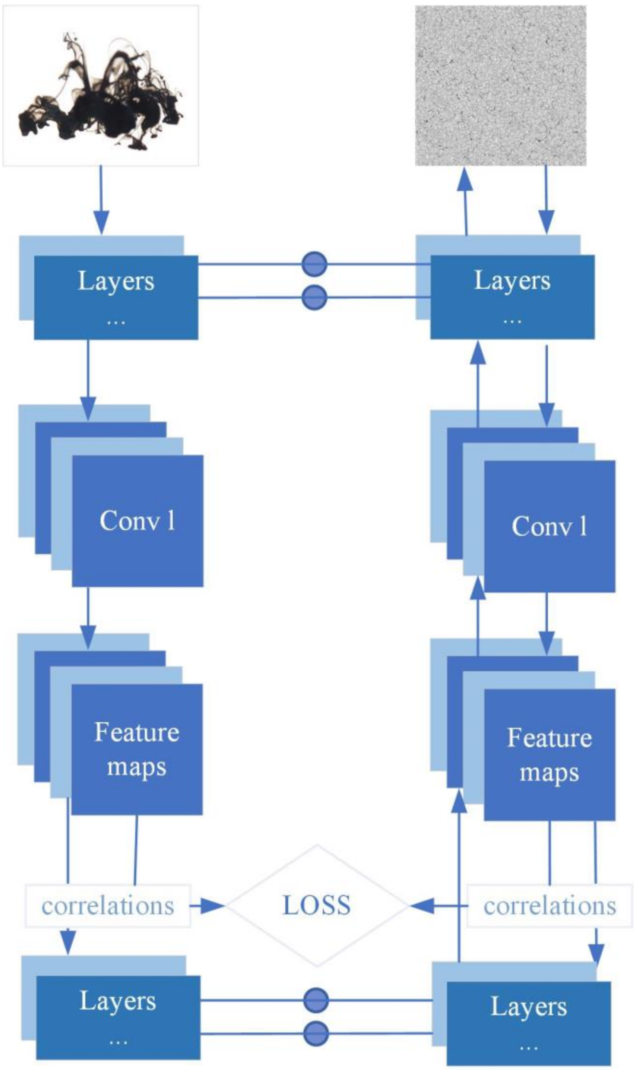

Figure 3: Extraction process of ink texture

To describe a given vectorized texture $\vec{x}$ in the texture model, it has to calculate its activation value for each layer l in the network through CNN. Since each layer in the network can be understood as a non-

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

linear filter bank, the image corresponding to the activation value constitutes a set of filtered images, which can also be called the features map. If there are N different filters in layer l, then there are N features graphs. If the size of the features graph is M, the features response in layer l can be expressed as a matrix:

$$F^l \in R^{N_l \times M_l} \tag{1}$$

In the above equation (1), F_{ik}^l represents a feature response of the i-th filter in the l-th layer at the k-th position of the features map corresponding to the i-th filter. Each texture has its fixed layout distribution, so the texture model is unknowable in terms of spatial information.

To obtain the ink texture information, the gradient descent method is used to match the Gram matrix of the given ink texture image on a white noise image [24]. This method can be achieved by minimizing the root mean square difference of the Gram matrix between the ink texture image and the white noise image. If $\vec{a}$ is set to be the vectorized source ink texture, Al represents the Gram matrix corresponding to layer l; $\vec{x}$ is the vectorized white noise image, and Gl represents the Gram matrix corresponding to layer l, then the loss function is expressed as below equation:

$$E_l = \frac{1}{4N_l^2 M_l^2}\sum_{i,j}\left(G_{ij}^l - A_{ij}^l\right)^2 \tag{2}$$

The total loss is expressed as equation (3):

$$L(\vec{a}，\vec{x}) = \sum_{l=0}^{L} W_l \, E_l \tag{3}$$

W in the equation above refers to the weight influence of each layer on the total loss. The derivative of Ffeatures response can be calculated based on equation (3):

$$\frac{\partial E_l}{\partial F_{ik}^l} = \left\{ \begin{array}{cc} \frac{1}{N_l^2 M_l^2}\left(\left(F^l - \frac{1}{(F^l+1)^2}\right)^T (G^l - A^l)\right)_{ki} & if\ F_{ik}^l > 0 \\ 0 & if\ F_{ik}^l < 0 \end{array} \right\} \tag{4}$$

According to the error backpropagation algorithm, the gradient of the loss function El and the gradient of $\boldsymbol{L}(\vec{a},\vec{x})$ with respect to $\vec{x}$ can be quickly obtained.

2.4 Ink photo synthesis based on CNN

The content features of the image include the appearance features (color, texture, and shape) and semantics of the image. The appearance features of the image belong to the low-level level, which is very intuitive to the observer. For high-level semantic features, observers are required to have certain subjective judgment and extraction capabilities. In fact, many image features extraction algorithms are compression of image content, and such a process is similar to human observation.

(1) Photo image enhancement based on Hue-Saturation-Value (HSV) color space

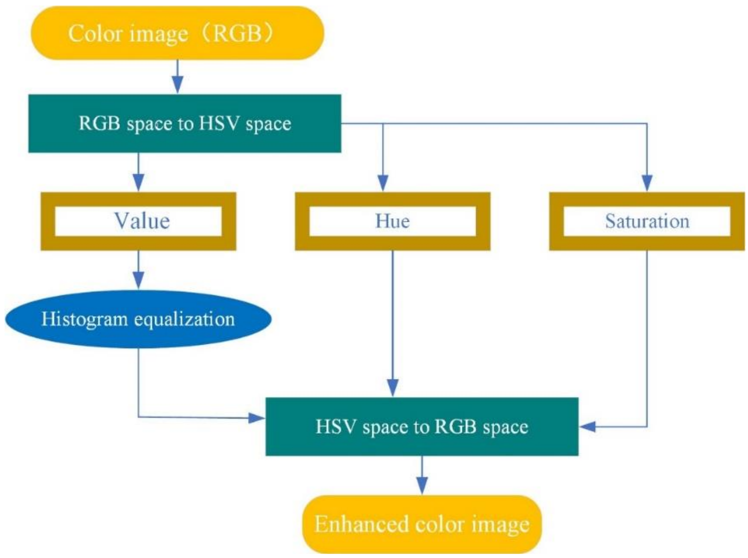

Figure 4: Image enhancement process.

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

Wetting is the diffusion effect of ink, which is one of the significant features of ink painting. When the wet brush touches the rice paper, the ink will flow from the brush into the rice paper and then infiltrate and penetrate the rice paper [25]. The CNN is adopted for image style extraction, considering the overall style of the image and weakening the details of texture. Therefore, it is correct when the ink painting is synthesized to make the generated image have a black-to-white ink diffusion effect on the boundaries of different grayscale areas. The input original photo image is enhanced to improve the contrast. In addition, it achieves the purpose of enhancing the content of the photo image. The method of image enhancement [26] is shown in the Figure 4:

The input photos are prone to color shift if they are directly processed based on the three components: red (R), green (G), and blue (B). Firstly, the image was converted from the RGB color space to the HSV color space. The three basic attributes of the HSV color space are H, S, and V. The HSV color model [27] is as shown Figure 5:

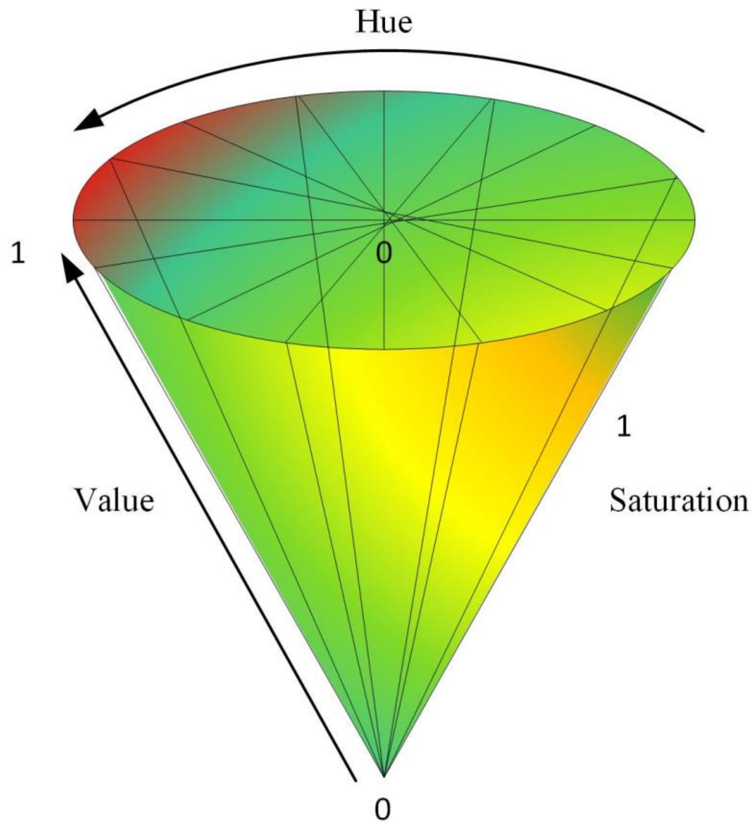

Figure 5: HSV model.

The conversion method is as follows:

$$\begin{cases} V = MAX \\ S = \begin{cases} 0 & if\ MAX = 0 \\ 1 - \frac{MIN}{MAX} & otherwise \end{cases} \\ H = \begin{cases} undefined & if\ MAX = MIN \\ 60 \times \frac{G-B}{MAX-MIN} & if\ (MAX = R) and (G \geq B) \\ 60 \times \left(\frac{G-B}{MAX-MIN} + 6\right) & if\ (MAX = R) and (G < B) \\ 60 \times \left(\frac{G-B}{MAX-MIN} + 2\right) & if\ MAX = G \\ 60 \times \left(\frac{G-B}{MAX-MIN} + 4\right) & if\ MAX = B \end{cases} \end{cases} \tag{5}$$

The value range of R, G, B in equation (5) is [0, 255]; the value range of S, V, and H is [0, 1], [0, 255], and [0, 360], respectively; MAX refers to the maximum of R, G, and B, and MIN represents the minimum of R, G, and B.

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

V is regarded as a grayscale image V1 for histogram equalization, and it is mapped into an image V2 with uniform grayscale distribution. The transformation function is the cumulative distribution function Sk of each grayscale probability:

$$S_k = T(r_k) = \sum_{j=0}^{k} \frac{n_j}{n} = \sum_{j=0}^{k} P_r(r_j) \qquad k=0,1,2,\ldots L\text{-}1 \tag{6}$$

In equation (6), rk is the normalized gray level of V1, nj is the number of pixels whose gray level is rj, n refers to the sum of pixels in image, T is the transformation relationship, and L represents the total number of possible gray level in image, Pr(rj) is the probability of the k-th gray value of the original image, and Sk is the normalized gray level of V2.

Finally, the image is converted back to the RGB color space, and the conversion is as follows:

$$\{(R,G,B) = \begin{cases} (v,z,x) & c_1 = 0 \\ (y,v,x) & c_1 = 1 \\ (x,v,z) & c_1 = 2 \\ (x,y,z) & c_1 = 3 \\ (z,x,v) & c_1 = 4 \\ (v,x,z) & c_1 = 5 \end{cases} \tag{7}$$

$$c_1 = \left[\frac{H}{60}\right] mod6 \tag{8}$$

$$c_2 = \frac{H}{60} - c_1 \tag{9}$$

$$\begin{cases} x = V \times (1 - S) \\ y = V \times (1 - S \times c_2) \\ z = V \times \big(1 - (S \times (1 - c_2))\big) \end{cases} \tag{10}$$

(2) Image content representation based on CNN

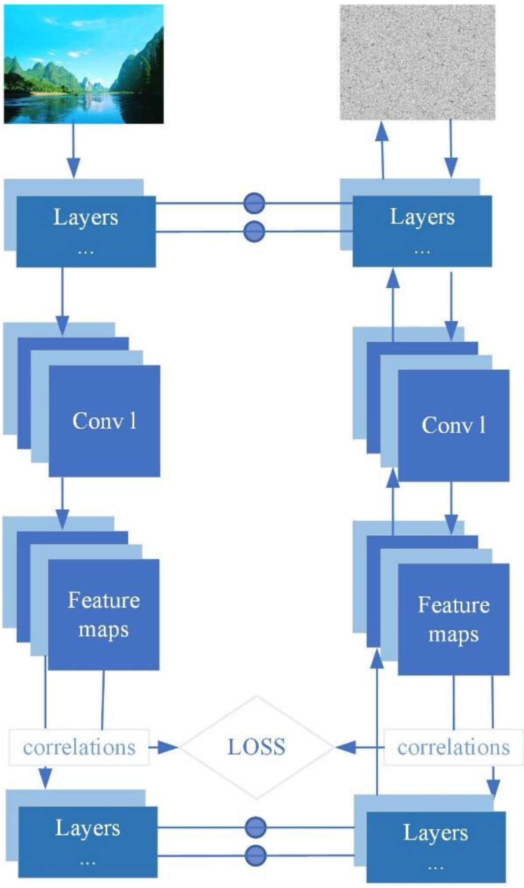

Figure 6: Image content conversion process.

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

CNN successively convolves the receptive field information of the input image through the features map (filter) [28] to obtain the image features information (features map) that is unchanged in translation, rotation and scaling, and then transmits the information to the higher layers in turn. When the CNN is trained for target recognition, the target information represented by the image features information becomes more and more clear as the number of layer increases [29]. The content of the photo is obtained by a random image matching the feature response of a photo in a certain layer of CNN [30]. The image content conversion process is shown in the Figure 6:

The gradient descent method is adopted to match the feature response of the original photo image on a certain layer of CNN on a white noise image, so that the white noise image is transformed into an image with the content information of the original photo. Let $\vec{x}$ and $\vec{p}$ be the vectorized photo image and white noise image, respectively, and P and F represent the corresponding features response matrix in layer l. The average error loss functions of P and F are defined as follows:

$$L_{content}(\vec{p},\vec{x},l)=\frac{1}{2}\sum_{i,k}\left(F_{ik}^{l}-P_{ik}^{l}\right)^{2} \tag{11}$$

The derivative of F can be solved:

$$\frac{\partial L_{content}}{\partial F_{ik}^{l}}=\begin{cases}(F^{l}-p^{l})_{ik} & if\ F_{ik}^{l}>0\\ 0 & if\ F_{ik}^{l}<0\end{cases} \tag{12}$$

In layer l, the gradient of $\vec{x}$ can be obtained through the error backpropagation algorithm, until $\vec{x}$ the feature response of the layer l mapped in this CNN is the same as that of $\vec{p}$ mapped in this layer, at this time $\vec{x}$ presents the content information related to $\vec{p}$.

(3) Generation of ink painting based on random image

To generate an image that mixes photo content and ink painting style, it has to perform the above-mentioned content acquisition and style acquisition operations on a white noise image at the same time. $\vec{p}$ is supposed to be a picture after image enhancement processing, and $\vec{a}$ and $\vec{x}$ are supposed to be an ink painting and a white noise image, respectively, then the loss function is expressed as equation (13):

$$L_{total}(\vec{p},\vec{a},\vec{x})=\alpha L_{content}(\vec{p},\vec{x})+\beta L_{style}(\vec{a},\vec{x}) \tag{13}$$

In equation (13), α and β are the influence weights of photo content and ink painting style in the generated image, respectively.

(4) Quick generation of ink painting based on content image

To quickly generate an ink image, it can directly obtain the style information of another ink image on the basis of a photo image. It is supposed that $\vec{p}$ is a photo after image enhancement processing, and $\vec{a}$ is to be an ink painting. The function can be expressed as below equation:

$$L_{total}(\vec{a},\vec{p})=\beta L_{style}(\vec{a},\vec{p}) \tag{14}$$

3. Results and Discussion

3.1 Synthetic effect of ink texture based on CNN

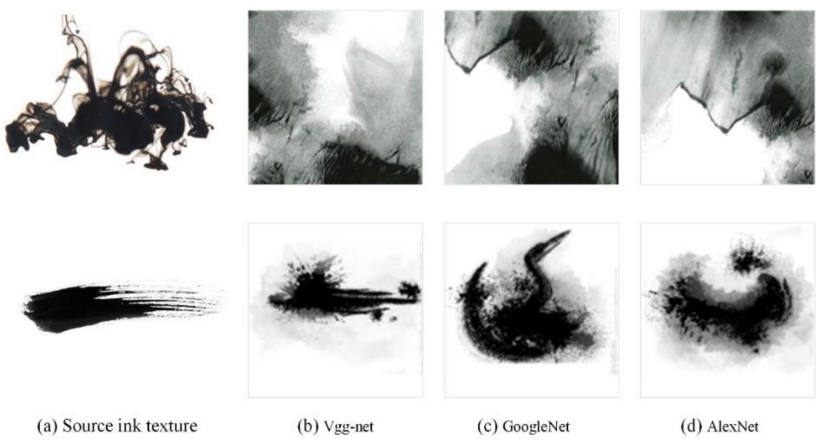

Figure 7: Comparison of ink texture synthesis effect.

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

Figure 7 (a) is the source texture image, and Figures 7 (b), 7(c), and 7(d) are the synthesized new texture images. The Figure 7 show that the effect of using the VGG-net network model is better than the other two. The convolutional layers used here are 'conv1_1', 'conv2_1', 'conv3_1', 'conv4_1', and 'conv5_1', that is, the weights (wl) of these convolution layers in equation (3) are all set to 1.

The feature summary statistical model based on the neural network model is used as the representation of ink texture, and a new ink texture is generated on a white noise image. This sample-based texture synthesis method has a better effect.

3.2 Comparison on photo image enhancement based on HSV and image content representation effect

(a) Source photos (b) Enhanced photos

Figure 8: Comparison of image enhancement processing results.

As shown in Figure 8, after the image is enhanced, the details of the image are more obvious, which is very helpful for extracting the texture later.

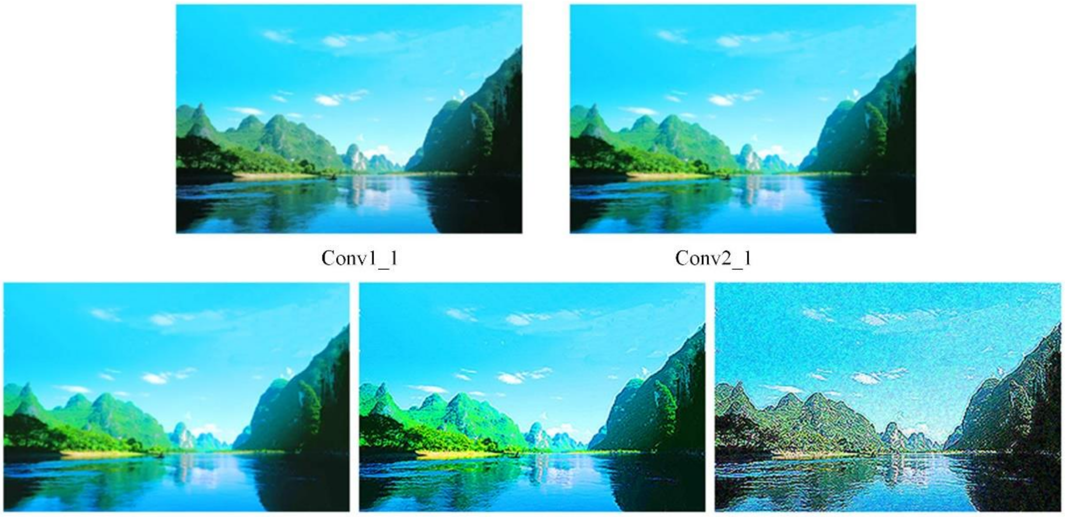

Figure 9: Comparison of image content generated by different convolutional layers of VGG-net.

In Figure 9, the content image generated by the low-level network shows relatively clear original photo information, while the high-level ignores the details of the image and mainly represents the content of the image. Although the image generated in the last layer lacks a lot of pixel information, it can still easily recognize this image because its main features information is preserved.

The content features of the image include the appearance features (color, texture, and shape) and semantics of the image. The appearance features of the image belong to the low-level level, which is very intuitive to the observer. For high-level semantic features, observers are required to make subjective judgment and extraction capabilities. It further shows that the image content representation effect synthesized by neural network is better.

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

3.3 Random image and image content as the basis to generate ink painting synthesis effect

The experimental parameters are as follows: the convolutional layer used to obtain the content of the photo is'conv4_2', and the convolutional layer used to obtain the ink painting style is 'conv1_1', 'conv2_1', 'conv3_1', 'conv4_1', and 'conv5_1'. In the equation (3), the weights (Wl) of these convolutional layers are all set to 1, and the others are all 0. In the total loss function equation (13) of ink painting synthesis, α is set to 1, β is set to 5, and the synthesis effect is shown in the Figure 10:

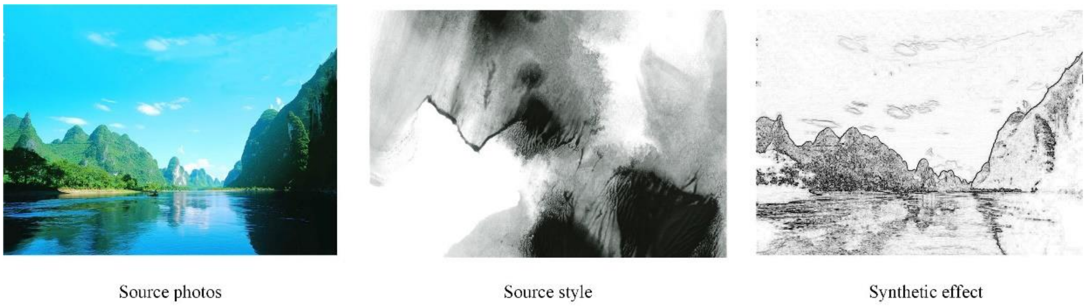

Figure 10: Random image as the basic composite effect display.

As shown in Figure 10, the ink painting effect synthesized on the basis of random image combines the characteristics of the two pictures. It not only retains the pixel effect of the original photo, but also integrates the ink texture extracted from the second picture, and the synthesis effect is better.

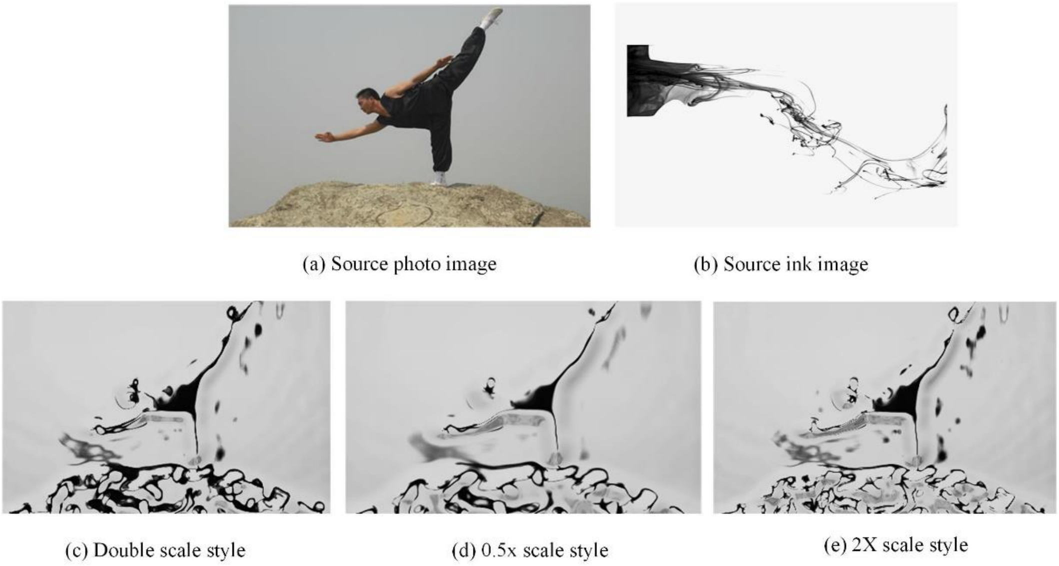

Figure 11: Synthetization of ink photos of different sizes and styles.

Figure 11 shows that it is feasible to synthesize different "size and style" ink photos. It shows that the ink painting synthesis based on CNN extends the possibility of computer creation of richer ink painting works.

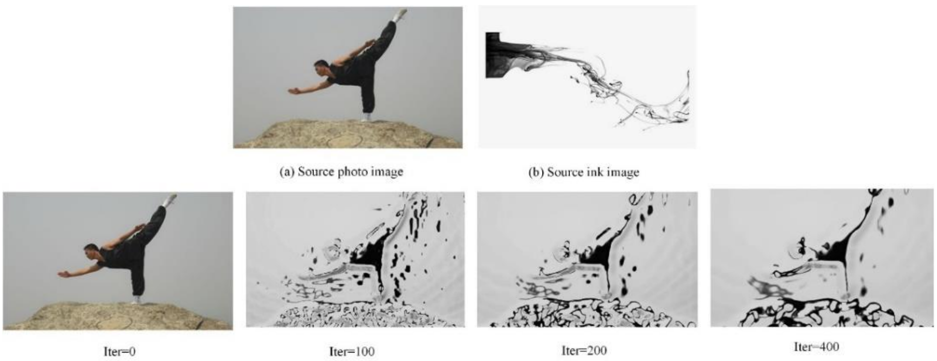

Figure 12: Quick synthetization of ink image effect.

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

As shown in Figure 12, Iter represents the number of iterations. The L-BFGS used by the optimization algorithm has a learning rate of 10. Although the ink image synthesized based on the content image has a slightly worse effect, it can quickly feel the rapid synthesis process.

4. Conclusion

The combination of computer science and art is a major direction of modern art, and the use of computers to simulate traditional painting art is even more popular. Chinese ink painting shows varied and rich ink techniques, forming a unique black and white art on rice paper. In the field of non-realistic sense, the simulation of ink painting is more difficult than the simulation of other painting art. In this article, a different ink image generation method is proposed from the past, which can generate ink images of different styles. The computer art simulation, NPR, especially some excellent ink painting drawing methods are investigated. The CNN model is studied, the basic attributes of image features are analyzed, and the basic characteristics of ink painting are considered to formulate the algorithm structure. The ink texture is synthesized based on the CNN model, which can be realized with following steps. The interrelationship of the features response calculated by the network is undertaken as the texture representation, and a white noise image is adopted to match the texture representation of the source texture image to generate a new ink texture image. The ink photo image based on CNN model can be synthesized with following operations. The texture of an ink image is synthesized to a photo with content information to generate ink photo image. According to the final experimental results, the ink image synthesized based on the CNN model has a better effect.

In this study, the relevant research on ink painting drawing is analyzed and reviewed, and a new drawing method is proposed aiming at the problem that the ink painting drawing method can only generate one or several fixed styles in the past. However, the method proposed in this study also has shortcomings and can be further studied. The ink synthesis method proposed in the study is based on the trained CNN model, and VGG-net is designed for image classification. Therefore, the VGG-net model selected in this study only selected based on the comparison of experimental results, so it is by no means the best feature extraction model. In the future, a CNN for image stylization can be specially designed.

References

[1] Sun D, Zhan R, Zhang J, et al. Realistic Story Visualization from Online Photo Collections [J]. Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao/Journal of Computer-Aided Design and Computer Graphics, 2018, 30(8):1479.

[2] Zhao Shaoyan. Style analysis of Wang Yuan's ink painting of flowers and birds [J]. Fine arts, 2019, 000 (002): 134-135.

[3] Zhang F, Gao H, Lai Y. Detail-Preserving CycleGAN-AdaIN Framework for Image-to-Ink Painting Translation [J]. IEEE Access, 2020, PP (99):1-1.

[4] Hauptfleisch F, Texler O, Texler A, et al. StyleProp: Real-time Example-based Stylization of 3D Models [J]. Computer Graphics Forum, 2020, 39(7):575-586.

[5] Huang L, Hou Z X, Zhao Y H, et al. Research progress on and prospects for virtual brush modeling in digital calligraphy and painting[J]. Frontiers of Information Technology & Electronic Engineering, 2019, 20(10):1307-1321.

[6] Zheng Y, Chen Y, Sarem M. A Novel Pencil Drawing Algorithm Based on Non-Symmetry and Anti-Packing Pattern Representation Model [J]. IEEE Access, 2019, 7:1-1.

[7] Kumar M, Poornima B, Nagendraswamy H S, et al. A comprehensive survey on non-photorealistic rendering and benchmark developments for image abstraction and stylization [J]. Iran Journal of Computer Science, 2019, 2(3):131-165.

[8] Montesdeoca S E, Seah H S, Semmo A, et al. Mnpr: A framework for real-time expressive non-photorealistic rendering of 3d computer graphics[C]//Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering. 2018: 1-11.

[9] Bakurov I, Ross B J. Non-photorealistic rendering with cartesian genetic programming using graphics processing units[C]//International Conference on Computational Intelligence in Music, Sound, Art and Design. Springer, Cham, 2018: 34-49.

[10] Qian W, Cao J, Xu D, et al. Research status and Prospect of nonphotorealistic rendering technology [J]. Chinese Journal of image graphics, 2020 (7): 1283-1295.

[11] Liu A, Zhao Z, Zhang C, et al. Smooth filtering identification based on convolutional neural

ISSN 2618-1568 Vol. 4, Issue 15: 95-107, DOI: 10.25236/FAR.2022.041519

networks[J]. Multimedia Tools and Applications, 2019, 78(19):26851-26865.
[12] Seddiki K, Saudemont P, Precioso F, et al. Cumulative learning enables convolutional neural network representations for small mass spectrometry data classification[J]. Nature Communications, 2020, 11(1).
[13] Tan M, Li J, Xu G, et al. A Novel Intuitionistic Fuzzy Inhibitor Arc Petri Net with Error Back Propagation Algorithm and Application in Fault Diagnosis [J]. IEEE Access, 2019, PP (99):1-1.
[14] Wang J, H Wu, Zhang X, et al. Watermarking in Deep Neural Networks via Error Back-propagation[J]. Electronic Imaging, 2020, 2020(4):22-1-22-9.
[15] Zeng J L, Chen K H, Wang J Y. High-Performance Vlsi Design For Convolution Layer Of Deep Learning Neural Networks[J]. Journal of electrical engineering, 2019, 26(5):195-202.
[16] Zhang T, Hu D, Zhang J. Scale Adaptive Block Target Tracking Based on Multi-layer Convolution Features and Kernel Correlation Filter[J]. Journal of Physics Conference Series, 2021, 1828(1):012042.
[17] Leclaire A, Rabin J. A Stochastic Multi-laryer Algorithm for Semi-discrete Optimal Transport with Applications to Texture Synthesis and Style Transfer [J]. Journal of Mathematical Imaging and Vision, 2021, 63(2):282-308.
[18] Chen Q, Li G, Xie L, et al. Structure guided image completion using texture synthesis and region segmentation[J]. Optik, 2019, 185:896-909.
[19] Al-Jaberi A K, Hameed E M. A Review of PDE Based Local Inpainting Methods [J]. Journal of Physics Conference Series, 2021, 1818(1):012149.
[20] Badawi A A, Chao J, Jie L, et al. The AlexNet Moment for Homomorphic Encryption: HCNN, the First Homomorphic CNN on Encrypted Data with GPUs [J]. IEEE Transactions on Emerging Topics in Computing, 2020, PP (99):1-1.
[21] Xie S, Zheng X, Chen Y, et al. Artifact Removal using Improved GoogLeNet for Sparse-view CT Reconstruction[J]. Scientific Reports, 2018, 8(1):6700.
[22] Amir S, Faturrahman, Hendra. Hand posture classification with convolutional neural networks on VGG-net Architecture [J]. IOP Conference Series: Earth and Environmental Science, 2020, 575(1):012186 (8pp).
[23] Sehad A, Chibani Y, Hedjam R, et al. Gabor filter-based texture for ancient degraded document image binarization [J]. Pattern Analysis and Applications, 2019, 22(1):1-22.
[24] Dai Y, Li Y, Sun B, et al. Skip-Connected Network with Gram Matrix for Product Image Retrieval [J]. Neurocomputing, 2021, 447(3).
[25] Deng s, Che X. GPU based 3D model rendering of Chinese ink painting [J]. Computer aided graphic design and manufacturing (English version), 2017 (1): 1-6.
[26] Cho S W, Na R B, Koo J H, et al. Semantic Segmentation with Low light Images by Modified CycleGAN-based Image Enhancement [J]. IEEE Access, 2020, PP (99):1-1.
[27] Zhang L, Technology of Low Contrast Printing Defects Based on Enhancement. Packaging Engineering, 2019, 40(13):252-258.
[28] Liang Y, Liu W, Yi S, et al. Filter pruning-based two-step feature map reconstruction [J]. Signal Image and Video Processing, 2021(2).
[29] Hababeh I, Mahameed I, Abdelhadi A, et al. Utilizing Convolutional Neural Networks for Image Classification and Securing Mobility of People with Physical and Mental Disabilities in Cloud Systems[J]. IEEE Access, 2020, PP (99):1-1.
[30] Xin Y, Jwab C, Wei H. Hybrid fuzzy integrated convolutional neural network (HFICNN) for similarity feature recognition problem in abnormal netflow detection [J]. Neurocomputing, 2020, 415:332-346.