Research on Chinese Character Style Transfer based on Convolutional Neural Network

Chang Cun, Xiaoxu Cui, Zhanhao Liang

[1] Hainan University, Institute of Tropical Agriculture and Forestry, Hai Kou, Hai Nan 252500, China

[2] Sichuan Agricultural University, Institute of Information Technology, Cheng Du, Cheng Du, 625000, China

[3] Guangdong Ocean University, Electronics and Information engineering, Zhanjiang, Guangdong, 524088

Abstract

Style migration is a new topic in the field of computer vision. The related algorithms of deep learning play a huge role in it. This paper introduces the image and Chinese character style migration algorithm based on convolutional neural network. The VGG-19 classifier is flexible. On the one hand, the picture style migration algorithm proposed in this paper can quickly generate a variety of different effects of the image, which can be applied to video post-processing, movie poster design and artistic creation. On the other hand, the Chinese character style migration algorithm proposed in this paper also saves a lot of time for the text experts to design different styles of Chinese fonts, which satisfies the actual needs of the font style contained in the Chinese font library to a certain extent.

Keywords

Convolutional neural network; Style migration; Chinese character style.

1. Introduction

In recent years, as deep networks have made breakthroughs in computer vision-related fields, some researchers have begun to focus on artistic creation, and style migration is a relatively successful attempt. Style migration is a fusion of classic art form and artificial intelligence technology, which has a great impact on both the art and technology fields. Chinese character style migration refers to the extraction of features from a specific Chinese character style for style conversion, and the conversion of ordinary Chinese characters into Chinese characters of ideal style.

2. Logic implementation

When generating the style feature texture representation vector, the first thing is to read the style image, and then forward the way to read the features in the deep convolution network respectively. stride shows that the filter moves horizontally and vertically in the picture. As 1, padding indicates that padding is added to the edge of the image, and the image after the convolution is kept the same size as the original image. Since the size of the convolution kernel is 3x3, the edge of the original image is filled with 1 layer 0, and the forward direction is The algorithm can be expressed as:

$$a^2 = \sigma\left(Z^2\right) = \sigma\left(a^1 + w^2 + b^2\right)$$

Among them, the superscript represents the number of layers, the asterisk represents the convolution, and b represents our bias, and σ is the activation function, which is generally ReLU. After passing the VGG-19's forward table, wait for its characteristic response, then use the Grim matrix for style description, because different convolutional layers represent different style features of the image, deep convolutional layers and shallow layers. The features selected by the convolutional layer need to be combined with each other to represent the style texture closest to the real image.

Through these, the texture representation of the style image can be obtained from the feature response, and the difference between the response values of the Gram matrix of the ink image and the noise image represents the similarity between the two images. An image having the same texture as the original image is gradually generated by the optimization algorithm. A represents the original image, WG represents the generated image of the noise image, N represents the number of feature maps of one layer, and M represents the size of the feature map. Then the loss of each layer can be normalized as:

$$E = \frac{1}{4N^2M^2} \sum_{\mathrm{i,\ j}} (G_{\mathrm{i,\ j}} - A_{\mathrm{ij}})^2$$

Since we choose multiple layers to represent the texture similarity of the image, the total loss of all layers can be W_i The influence factor of the loss value of each layer on the total loss value can be reduced by W_l To reduce the impact of this layer, if you want to force the effect of a layer, W The corresponding impact factor of this layer:

$$L_{style} = \sum_{l=0}^{L} w_l E_l$$

After calculating the loss function of the style picture, the pixel value of the noise picture is iteratively updated by the gradient descent algorithm, and moved to the direction in which the loss function value decreases. The pixel value update of the noise picture is as shown in the formula, where Mow' indicates the updated pixel value, Mow indicates the pixel value before the update, and II is the set learning rate, which can be obtained by the chain derivation rule:

$$x^{'} = x - \mu \frac{\partial L}{\partial x}$$

The final loss function form is:

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

among them $\overrightarrow{p}$ For the content image, α It is a style image. α with β Is the weight coefficient of the content image and the style image, the ratio of the two α / β Deciding who to generate the image is more affected.

3. Content feature extraction

In order to make the generator better realize the style migration, it is necessary to analyze and extract the content features of the Chinese characters in the standard sample image. Considering that Chinese characters are composed of basic strokes such as horizontal, vertical, Left, Na, dots, and folds, a Chinese character can be determined by the shape parameters and coordinates of strokes. Therefore, the extraction of Chinese character content can be converted into the characteristics of strokes. extract. The feature description parameters of strokes are not unique. This paper mainly selects the following five parameters to form Chinese characters. $Word$ Feature set:

$$Word_i = \{g_i, se_i, l_i, w_i, k_i\} \qquad i = 1, 2, ..., n$$

among them, n Chinese characters $Word$ The total number of strokes in the middle, $Word_i$ For the first i a set of features of strokes, g_i a set of features for each skeleton point for the stroke, se_i a collection of starting points (including start and end points) for the stroke, l_i For stroke length, w_i For stroke width, k_i The key point for strokes.

Based on the feature set of Chinese characters, the Chinese character image of the standard sample is processed:

*S*tep1: The Chinese image is normalized, and the size of the unified image is 160×160 Pixel

*S*tep2: Image binarization, using the first open and then closed operations for denoising, reducing the amount of calculation;

*S*tep3: Refine the image to obtain the skeleton feature of the Chinese character;

*S*tep4: Extract key points: According to the eight neighborhood search method, the endpoints and inflection points of the skeleton are extracted. When only one point in the eight neighborhoods of a skeleton point is black, the change point is considered to be the skeleton end point, and the two end points are directly skeleton points. The number of strokes is the length of the stroke;

*S*tep5: Extract the stroke width: take the initial skeleton point pixel coordinate as the center of the circle, the initial radius is r to draw the circle, check whether each pixel in the circle is within the stroke range, that is, check whether there is a white point in the circle, and if there is no white, follow a certain step. Long expansion r continues to judge until there is a white point in the circle, taking 2(r-1) at this time as the stroke width at the current skeleton point, and taking the next skeleton point as the center of the circle repeats the above steps.

4. Style feature extraction

Compared with the RGB three-dimensional color image in the picture style migration, the image of the Chinese character is two-dimensional, and the process of convolution is relatively simple. Convolutional neural network with the deepening of the number of network layers, the extracted features are more and more abstract, more able to reflect the style characteristics of the target sample.

Input layer: each Chinese character image in the training set is 160×160 Pixel, that is, the image is first read 160×160 The matrix acts as an input to the convolutional neural network.

Convolutional layer: The main purpose of convolving the input matrix is to extract the feature information of the image. The convolution kernel is a 8×8 a matrix having a step size of 2, convolving the input image matrix by one pixel at a time, multiplying the different local matrix of the input image and the elements of each position of the convolution kernel matrix, and adding the result matrix in the convolution The value of the location, such that the feature extraction of the convolutional neural network can preserve its location characteristics.

Pooling layer: Compared to the complexity of the convolutional layer, the pooling layer is much simpler, that is, compressing each sub-matrix of the input tensor. The common pooling standard is MAX and Average, that is, the maximum or average value of the corresponding region is taken as the pooled element value. Here we use the more commonly used MAX pooling method, while using 3×3 Pooling, the stride is 2.

Fully connected layer: Use the ReLu activation function in the output layer of the fully connected layer to get 80×80 Output matrix.

5. Style migration

In the step of generating an image in the Chinese character style migration, using the feature extracted by the above convolutional neural network as a part of the objective function, comparing the stroke feature set of the standard sample and the feature matrix output by the convolution operation of the target sample, using the Deep Dream algorithm Combine the two to generate the target Chinese characters.

The final synthesis results are shown below:

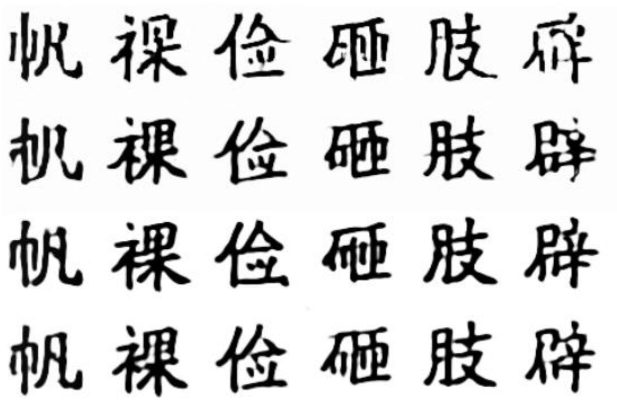

Figure 1 Chinese character synthesis results

6. Conclusion

In view of the outstanding performance of deep learning algorithms such as convolutional neural networks in image feature extraction, and the convergence of VGG-19 network with style extraction and content extraction tasks, convolutional neural networks have become the mainstream framework for major research in this field. The proposed style migration algorithm means that style-based migration can be applied in more aspects, and the business prospects are very clear. In future research, we will work to overcome the shortcomings of convolutional neural networks and apply style migration in more ways.

References

[1] Li Chaoqi, Wang Shaoyu.Face age classification based on deep convolutional neural network[J]. Intelligent Computer and Application. 2017 (03).

[2] Dong Wei, Ding Youdong, Qian Wei.Application of Artificial Intelligence Based Style Migration Algorithm in Animation Special Effect Design[J]. Decoration. 2018 (01).