

An emotional expression framework based on social media style transfer

Chuying Zhou

Mathematics and Systems Sciences, Guangdong Polytechnic Normal University,
Guangzhou, 510180, China

chuyingzhou2378@163.com

Abstract. Recently, artificial intelligence has great effort in social media. Style transfer is a subtask of deep learning, can transform images shared on social media into works with artistic painting style transfer through image recognition technology, which creating amazing effects. However, the current style transfer technology is still in the early stages of development, and can only convert the shared picture into the style of a certain artwork, but does not have the ability to convert the emotional expression of the picture. In this paper, we proposes an emotional expression framework based on social media style transfer, conducted an experiment in the LFW database where the same portrait photo can express different emotions by combining it with images of different styles and realizes the style transfer technology for emotion and emotional expression. Therefore, the application of style transfer technology in social media is of great value and can promote the intelligent development of social media.

Keywords: artificial-intelligence, style-transfer, portrait-emotion, social-media.

1. Introduction

1.1. Good development prospects for artificial intelligence

With the development of artificial intelligence (AI), various industries have promoted the upgrading of industries through AI, and various countries and regions have expressed high hopes and strong support for the development of AI industry, and continue to land on the relevant emerging applications and policies. As the chart below shows, the global AI industry has grown from \$690 billion five years ago to a 2021 of \$3 trillion, and is expected to surpass \$6 trillion by 2025. As shown in Figure 1, from 2017 to 2025, it is expected to grow rapidly at a compound growth rate of over 30%.

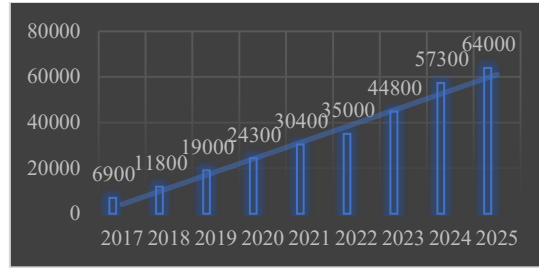


Figure 1. Global artificial intelligence industry scale from 2017 to 2025 (unit: US\$ billion).

1.2. One of the applications of artificial intelligence in the image field

With the development of artificial intelligence, its influence is more and more profound in the application of visual language. For example, there are image classification and style transfer in image domain. As you can see, style migration is an example of AI in visual language applications, and it has the potential to be widely used in social media. As shown in Figure 2, the main function of the computer is based on specific content automatically generate a specific style of image.



Figure 2. Schematic diagram of style transfer.

1.3. Transportation sustainability design strategies for general commercial complexes

Neural network-based image style transfer emerged, which was previously unimaginable for a program to imitate any image, and it was presented in 2015 by Gatys et al. in two papers: Gatys et al., 2015a and Gatys et al., 2015b, which gave a way to model textures using deep learning. As shown in Figure 3, the rendering of the image style transfer technology based on the neural network.



Figure 3. Rendering of Image Style Transfer Technology Based on Neural Network.

1.4. Transportation sustainability design strategies for general commercial complexes

However, the traditional image style transfer method is based on pixel level, work style, and we are based on deep learning to extract features, no matter what genre author, the shortcomings of traditional methods, style and content weight can not be properly controlled, learning ability and expansion effect is limited. Therefore, this paper proposes a method based on deep learning, which trains people through limited real photos of people to ensure that people can use fewer photos to vividly convey different emotions in social media, promote characters to use their own photos to communicate emotions in social media, and help characters better express emotions. A theoretical system was made for the subsequent photo style transfer system in social media. For example, select a photo of a person, and convert the picture into a picture with the same content that can express obviously sad and hurt emotions according to the sad emotional style set by the system in the social software. As shown in Figure 4, the photo (a) is the original picture, the photo (b) after style conversion feels bad, unfriendly, and has nothing to do with happiness, and the photo (c) after another style conversion to the right picture, obviously bright and happy. That's where image style migration comes into a role in social software—expressing more emotions with a limited number of images of people. After comparing the application effects of vgg16 (Visual Geometry Group 16) and vgg19 (Visual Geometry Group 19) in social media style transfer, it was decided to use vgg19 network structure, and the image results of style transfer were more in line with the user's content in social media[1].

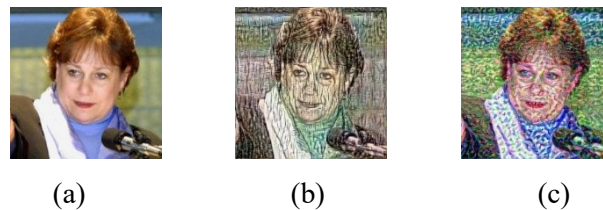


Figure 4 Style Conversion Diagram.

2. Case study

2.1. Neural network

Neural network is an important part of deep learning, which determines the training performance of the model, and the VGG (Visual Geometry Group, VGG) of the University of Oxford proposed the VGG16 model to solve the problem of accuracy, including convolution kernel becoming smaller and network deepening. Based on the decrease in accuracy of the training set, author He Kaiming proposed Identity Shortcut Connection, which directly skips one or more layers, which solves the problem of backpropagation gradient diffusion caused by the deepening of the number of network layers.

2.2. Computer vision

The Science of Computer Vision is the study of how to make a machine “See” like a human, or, more specifically, combine a computer with a camera, the machine vision related operations such as recognizing, tracking and measuring the target like the human eye, and further doing graphics processing[2]. As a scientific discipline, the leading scholars in the field of computer vision try to build artificial intelligence systems that can get 'information' from images.

2.3. Style transfer

Style transfer is also part of deep learning in artificial intelligence [3]. Style transfer is a new technology that has begun to develop in recent years and is one of the most amazing applications of artificial intelligence in creative environments, through image recognition and style transfer technology we are able to convert the selected image into the desired artistic painting style, thus creating amazing effects.

3. Method

3.1. Structure diagram of convolutional neural networks(vgg19)as shown in Figure 5

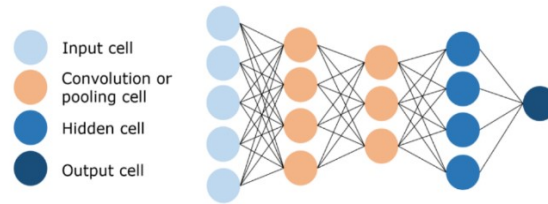


Figure 5. Structure diagram of convolutional neural networks(vgg19).

Convolutional neural networks (CNN) are a widely used type of artificial neural networks, especially in the fields of computer vision and natural language processing. It is capable of processing input data with two- or three-dimensional shapes and has excellent performance in processing images, audio, text, etc. vgg19 is a CNN with a depth of 19 layers. Among them, it should be noted that the reason for choosing vgg19 is that the more layers of the neural network, the deeper the level of abstraction of the input features, and the higher the accuracy of its understanding.

3.2. The choice of optimizer

The activation functions Vgg16 and 19 both use the Relu(Rectified Linear Unit) function

Choice of optimizer for image updates: Adadelta, RMSprop, Adam are relatively similar algorithms that perform similarly in similar situations[4]. Where you want to use RMSprop with momentum, or Adam, you can mostly use Nadam to achieve better results. After the comparison, choose to use NadamL earning rate: To optimize, we use a weighted combination of two losses to obtain the total loss:

$$\text{style_weight}=1e-2$$

$$\text{content_weight}=1e4$$

Reduce the weight of these high-frequency errors by regularizing the high-frequency components of the image:30.

3.3. Activation function

Formulas for the Sigmoid function(The curves for Sigmoid functions as shown in Figure 6):

$$f(x) = \frac{1}{1+e^{-x}} \quad (1)$$

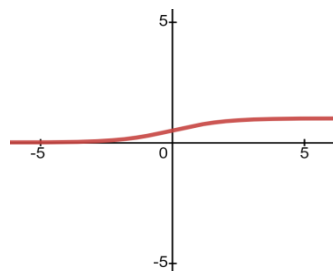


Figure 6. The curves for Sigmoid functions.

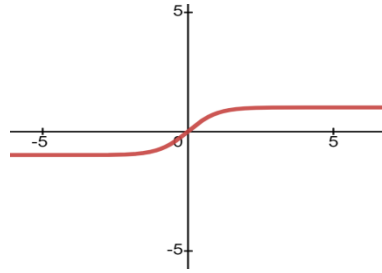


Figure 7. The curves for the Tanh function.

Formulas for Tanh functions(The curves for Sigmoid functions as shown in Figure 7):

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

The ReLU (Rectified Linear Unit) function is one of the most widely used activation functions. The formula of the Relu function are as follows(The curves for Sigmoid functions as shown in Figure 8):

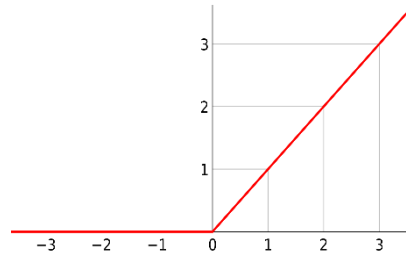


Figure 8. The curve of the Relu function [5].

$$ReLU(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \quad (3)$$

3.4. Description of my own method

The style and content of the image is extracted using a VGG19 convolutional neural network containing 16 convolution layers and 3 fully connected layers [6]. As shown in Figure 9, the VGG19 convolutional neural network is used in image style and content extraction, where each layer of the neural network takes advantage of the output of the next layer to extract further. In this process, more and more complex features are obtained until they can be used to recognize objects, so each layer of neural network can be regarded as a machine to extract local features. VGG19 consists of 16 convolution layers and the last 3 fully connected layers, as shown below, without the activation layer and pooling layer. The middle is the same as usual, with the pool layer, and finally through Softmax.

First, the image is input (None, 256, 256, 3), after passing through the convolutional layer of Block1_conv1 layers, the depth becomes 64, that is, the current format is (None, 256, 256, 64), and then pooled after a convolution Block1_conv2 layer - block1_pool layer, output (None, 128, 128, 64).

Then after passing through Block2_conv1 layers of convolutional layers, output (None, 128, 128, 128), and pooling after another convolution Block2_conv2 layer - block2_pool layers, output (None, 64, 64, 128).

Then after passing through Block3_conv1 layers of convolutional layers, output (None, 64, 64, 256), and then after Block3_conv2 layers, Block3_conv3 layers, Block3_conv4 layers of convolution, after pooling in block3_pool layers, output (None, 32, 32, 256).

Then after passing through the convolutional layer of Block4_conv1 layers, the output (None, 32, 32, 512), and then after the Block4_conv2 layer, Block4_conv3 layer, and Block4_conv4 layer three convolutional layers, after pooling in the block4_pool layer, the output (None, 16, 16, 512).

Then after passing through the convolutional layer of Block5_conv1 layers, output (None, 16, 16, 512), and then after passing through the convolutional layers of Block5_conv2 layer, Block5_conv3 layer, and Block5_conv4 layer, after pooling in the block5_pool layer, the output (None,8,8,512).

Finally, after three fully connected layers, softmax results are obtained.

VGG PROS: 1. The architecture of VGGNet is very simple. The convolution cores used throughout the network are of the same size (3x3), and the pool size is of the same size (2x2). 2. A convolutional layer composed of small filters is better than a convolutional layer composed of large filters.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 9. The network structure of vgg [7].

The activation function is relu function. It is not difficult to see from the above that the ReLU function is a piecewise linear function in which all negative values in ReLU become zero, and the positive value is still positive itself. This rule, called unilateral inhibition, makes neurons sparsely activated. For example, in a deep neural network model (such as CNN), when the model adds n layers, the activation rate of ReLU neurons will theoretically decrease by 2 to the nth power.

The role of neuronal sparsity used in this paper is that when training a deep classification model, there are often only a few features related to the target, so the model after sparseness is implemented through ReLU can better mine relevant features and fit the training data.

Based on the style and content of the extract, the new image update optimizer uses contrast before choosing to use NADAM. In essence, Adam (Adaptive Moment Estimation) is RMSPROP with momentum term, because it can dynamically adjust the learning rate of each parameter by using the gradient's first-order Moment Estimation and second-order Moment Estimation. The main advantage of Adam over RMSPROP is that it has an accurate learning rate for each iteration, which makes the optimizer's parameter selection process smoother. Nadam is similar to Adam with Nesterov momentum term. Nadam has a stronger constraint on learning rate and a more direct effect on the updating of gradients. The results show that NADAM can be used instead of RMSprop and Adam to achieve better results.

4. Experience

Here we are using the LFW database, Labeled Faces in the Wild, a database of face photos designed to study unconstrained face recognition problems. Compare different styles, as shown in Figure 10, select three styles (emotions) (sad, happy, neutral), and also perform deep learning processing on a portrait. Describe the resulting visualization in words.

Style:

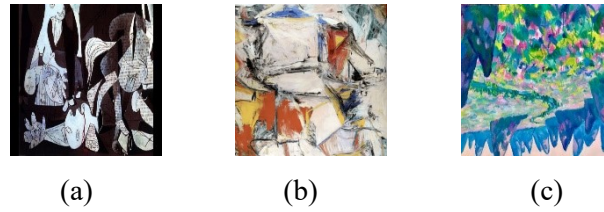


Figure 10. Emotional style schematics.(a) is sad, (b) is happy, (c) is neutral.

Content:

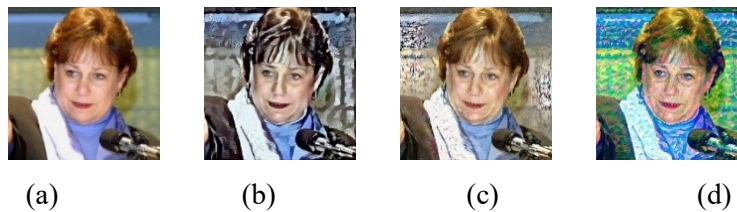


Figure 11. Schematic diagram of mood shift.(a) is Original, (b) is sad, (c) is happy, (d) is neutral.

As shown in Figure 11, now three different stylistic transfers of a photo (a). The comparison shows the representation of different emotional transfers of the same image of President XX in social media. Among them, the photo (b) with depressive emotional characteristics, resulting in a paler portrait and a lower overall atmosphere, contrary to an optimistic and positive state; For the photo (c), the overall tone is a little warmer than the emo style, and the mood is a little softer, but there is no particularly positive or negative attitude; Finally, the photo (d) is generated, the whole tone is more vivid, the character as a whole reflects the characteristics of happiness and relaxation, and the character reflects the breath of life. By comparing the results of repeated experiments, we obtain the parameters and network configuration that are most suitable for the emotional transfer technology of character images, and also verify the feasibility and creativity of integrating emotional transfer into style transfer technology.

5. Conclusion

As a common artificial intelligence deep learning technology, image style transfer is creative and exploitative. The traditional method is to get the style of the work through image transfer and pixel-level feature extraction, but the proper weight and feature representation can not be learned between the style and the content, which leads to the limited learning ability and generalization effect of the model, difficult to promote on social media. The solution to the above challenge is to design a learning-based model by choosing VGG19 as a convolutional neural network, comparing different optimizers. The model extracts image-style features and trains them with specific portraits so that publishers can convey different emotions with fewer photos on social media. We proposes an emotional expression framework based on social media style transfer, conducted an experiment in the LFW database where the same portrait photo can express different emotions by combining it with images of different styles and realizes the style transfer technology for emotion and emotional expression. In figures 10 and 11, the process of style transfer using the same portrait (photo(a), in Figure 11) to express the three different emotions in figure 10 is shown, resulting in the other three figures in Figure 11. It can be seen that the experimental performance is accurate and significant, which provides a train of thought for art analysis and provides more abundant results for the realization of the art field of AIGC, this paper constructs a theoretical system for the following portrait style transfer system in social media. Therefore, the application of style transfer technology in social media is of great value and can promote the intelligent development of social media.

References

- [1] LIU Minghao. Image style transfer based on VGG-16[J]. Electronic Production, 2020(12):52-54. DOI:10.3969/j.issn.1006-5059.2020.12.018.
- [2] Liu Zhe On Computer Vision Technology [J] Digital users, 2019,25 (8): 159. DOI: 10.3969/j.issn.1009-0843.2019. 08.149.
- [3] TANG Zhiwei, LIU Qihe, TAN Hao. Review of neural style transfer models[J]. Computer Engineering and Applications, 2021, 57(19):32-43. DOI:10.3778/j.issn.1002-8331.2105-0296.
- [4] Jiang Angbo, Wang Weiwei Optimization of ReLU activation function [J] Sensors and Microsystems, 2018, 37 (2): 50-52 DOI:10.13873/J.1000—9787(2018)02—0050—03.
- [5] CSDN. (2021) In-depth understanding of ReLU functions (Interpretability of ReLU functions). https://blog.csdn.net/weixin_41929524/article/details/112253138
- [6] Liang Jiazhi. Image Classification Based on VGGNET[J]. Digital User, 2019, 25(46):87-88.
- [7] CSDN. (2023) Classical Network Architecture Learning -VGG. <https://blog.csdn.net/BXD1314/article/details/125781929>