International Journal of Communication Networks and Information Security
2024, 16(1), 6117
ISSN: 2073-607X,2076-0930
https://ijcnis.org/
Research Article

Graphic Style Transfer Technology in Multimedia Communication: An Application of Deep Residual Adaptive Networks in Graphic Design

Zhenyu Zhang [1, 2], Fuyu Wei [3, 4*], Guizhen Liang [5], Xintong Wang [6]

[1] Ph.D., City Graduate School, Faculty of Creative Industries (FCI), City University Malaysia, Kuala Lumpur, Malaysia
[2] Lecturer, School of Fine Arts, Xinxiang University, Xinxiang, China
[3] Teaching assistant, College of Materials Sciences and Engineering, Henan Institute of Technology, No.699, Plain Road, Hongqi District, Xinxiang City, China
[4] Ph.D. Candidate, College of Computing, Informatics and Mathematics, Al-Khawarizmi Building, Universiti Teknologi MARA, 40450, Shah Alam Selangor Darul Ehsan, Malaysia
[5] Professor, School of Mathematics and Statistics, Xinxiang University, Xinxiang, China
[6] Master, School of Mathematical Sciences, Henan University of Science and Technology, Xinxiang, China

*Corresponding Author: weifuyucc@163.com

ARTICLE INFO

Received: 15 November 2023
Accepted: 6 April 2024

ABSTRACT

With the rapid development of wireless network technology and the rapid popularity of portable smart terminals, multimedia communication based on images and videos has become the favorite way of communication in the new era. Image style transfer technology is one of the research directions that has attracted much attention in the field of multimedia communication. To achieve the diversification of images and ease of use in the multimedia communication process, this paper researches the multimedia network communication technology and image style transfer technology. By combining visual style transfer technology and depth residual adaptive network technology in multimedia communication technology, the redesign and creation of graphics can be carried out effectively. The resulting graphics can meet the needs of the art creators and the technique provides higher creative efficiency, excellent peak model signal-to-noise ratio and structural similarity performance, and output levels that meet the basic needs compared to traditional manual design. The method can be effectively used in urban building appearance design and art creation and has good theoretical and practical research value.

Keywords: Multimedia Communication, Communication, Deep Learning, Art.

Citation: Z. Zhang, F. Wei, G. Liang, and X. Wang, "Graphic Style Transfer Technology in Multimedia Communication: An Application of Deep Residual Adaptive Networks in Graphic Design," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 16, no. 1, pp. 198-210, Apr. 2024.

INTRODUCTION

With the rapid development of wireless network technology and the rapid popularity of portable smart terminals, compared with traditional text-to-speech communication, image and video-based multimedia communication has rapidly become the most preferred mode of communication in the new era because of its more intuitive and closer to people's distance. With the rapid development of various live video sharing platforms and multimedia social platforms such as Beili Beili, YouTube, Jittery, WeChat, Skype, Twitter and other new multimedia applications, multimedia services such as online live streaming, real-time dynamic sharing, live video broadcasting, video conferencing, and other multimedia services based on uplink data streams have rapidly penetrated every aspect of people's daily lives and become the most popular means of communication in the new era [1]. It has become the main factor driving the explosive growth of network data traffic in the era of mobile Internet. With the gradual improvement of 5G network coverage, the development of the Internet of Things will be closer to and integrated into people's real life, low-power monitoring, remote area sensing, and other wireless multimedia sensor networks (facilities) will be more widely spread and covered. And behind these emerging multimedia applications, gradually emerging is a new type of multimedia application scenario. Both portable mobile intelligent terminals and wireless sensor network facilities on the signal acquisition side are faced with resource constraints, such as limited computing power, battery capacity, and storage space, etc., especially the problem of computing power and battery capacity has become a bottleneck limiting the further development of mobile intelligent terminals. Even for scenarios such as multimedia sensor networks, which do not have portability requirements, pure hardware stacking will lead to a drastic increase in network deployment costs, limiting the commercial development of multimedia services. Therefore, in the face of resource-constrained multimedia application scenarios, it is required that the signal acquisition side should maintain a low computational complexity, i.e., better support for new multimedia services should be achieved by reducing the resource consumption rather than upgrading the hardware facilities without any limitation [2].

The development of multimedia communication technology and deep residual adaptive network technology has led to a wide range of applications of image as an information carrier, and the increasing degree of data network connectivity has also increased the frequency of application of image information in social media [3]. As an essential medium for information transmission and a direct manifestation of artistic creation, images have become the primary medium and way for people to communicate technology in their daily lives [4]. When art creators create artistic graphics, not only do they gradually increase their requirements for mobile hardware devices, but their excellent shooting functions, standardization of images, and realistic rendering have become essential requirements for art creation. At the same time, people have also raised higher demands for the level and form of image processing, and their needs are no longer only focused on smoothing, restoration, enhancement, and other processing forms; instead, we look forward to providing personalized and artistic visual experiences for images [5], [6]. Therefore, research on multi-style beautification and artistic creation of images has become increasingly popular. In addition to the long-studied non-realistic rendering techniques [7], [8], the transfer technology of image style has also received widespread attention in recent years [9], [10], [11], [12].

Moreover, with the continuous updating of image processing software functions, people not only hope for unrestricted migration of new artistic styles to photos but also hope to highlight the foreground targets of interest in the images. Thus, the foreground target of the image is extracted, and the style is shifted, making the image both artistic and personalized. The emergence of image style transfer technology not only enhances visual aesthetics but also dramatically enriches people's cultural life [13]. Under the guidance and promotion of information technology, more and more animation and film production industries have utilized image-style transfer technology to achieve brilliant animation effects.

In recent years, with the innovation and replacement of technology and the improvement of computer computing power, neural networks have been developed rapidly, and their application in all walks of life has been gradually widespread, including the application of graphic style migration. This technology also has a significant impact on image style transfer and slowly produces two categories in Figure 1, namely, style transfer without neural network and style transfer with neural network. Style transfer without a neural network generally refers to the image art stylization realized by three methods, namely, stroke-based rendering [14], image category [15], [16] and image filtering [17] before the emergence of the neural network, which can be attributed to early image style rendering, that is, image style transfer without neural network [18], This is mainly the image style migration method based on convolutional neural network texture synthesis proposed by J. Zhou et al. [19], [20], [21], which separates and extracts the image content abstract features and style abstract features respectively, and then processes these high-level abstract features through the pre-training VGg model [22], [23], synthesizing an artistic effect image with original content and new style texture in an iterative optimization way, It effectively realizes the artistic effect of image style transfer. The neural network is an image-style migration method based on the cyclic generative countermeasure network proposed by S. K. Gorti et al. [24], [25], [26]. This method connects two one-way propagation generative countermeasure networks in a ring and solves the problem that image features need to rely on training data pairing when migrating in this ring structure network. The application that realizes the effect of image style migration refers to the application of neural network technology to learn image style migration technology, which mainly includes two kinds: one is online image optimization, and the other is the optimization application through the offline model.

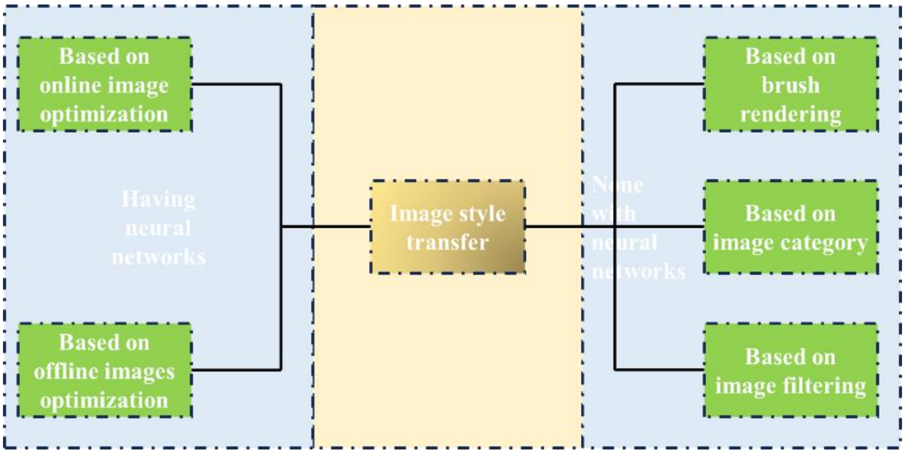

Figure 1. Classification of Image Style Transfer Methods

To realize the diversity and ease of use of images in the multimedia communication process, this paper investigates multimedia network communication technology and image style transfer technology. We hope to learn more ways of describing styles through computer training, using deep residual adaptive network models, and solving the problem of generating images with more desirable stylization effects and multi-domain migration by starting from deep learning-related techniques, to meet the requirements of people on images and image editing in multimedia communication process, which has obvious research significance.

LITERATURE REVIEW

Image Processing Techniques in Multimedia Communication Processes

The main aspects involved in multimedia communication technology can be categorized into data compression and coding technology, multimedia network technology, and so on.

Data Compression and Coding Technology

Data compression and coding technology has been the key technology in multimedia communication, image compression and coding from the beginning of systematic research so far has nearly 50 years of history, in 1988, CCITT formulated the visual telephone, conference television standards, and later MPEG-1 and other image compression standards are based on it. Compression coding techniques for multimedia data are developed based on C.E. Shannon's information theory. The coding methods can be divided into three categories: (1) Predictive coding, transform coding, vector quantization coding, sub-band coding, neural network coding, etc., according to the statistical characteristics of the information source (first-generation coding methods). (2) According to the visual characteristics of the human eye, image coding is based on directional filtering, image contour one-ethic coding, wavelet analysis-based coding, and other methods (second-generation coding methods). (3) According to the transmission of scenic characteristics: fractal coding, model-based coding, and other methods (second-generation coding methods) [27].

Multimedia Network Technology

Network multimedia technology is a kind of comprehensive technology integrating many kinds of technologies, using computer technology, network communication technology, and information processing technology to realize the collection, processing, storage, transmission, and display of text, picture, audio, video, and other media information. With the support of network multimedia technology, people can acquire, process, and exchange multimedia information more conveniently, thus greatly enriching the content and form of information. Network multimedia technology is built based on digital technology, through the conversion of various media information into digital signals to achieve rapid processing and transmission of information.

With the help of network multimedia technology, people can acquire, process, and share multimedia information more efficiently, further improving the efficiency and convenience of information exchange. In addition to digital technology, network multimedia technology also involves the integration and application of a variety of technologies, including audio technology, video technology, compression technology, storage technology, and so on [28]. The application of these technologies makes the collection, processing, and transmission of multimedia information more efficient and reliable, and at the same time provides more diverse forms and effects for the display of multimedia information.

Overall Framework Design of Multimedia Communication Systems

The multimedia communication system is based on the telephone, which adds the image transmission and display function based on the ordinary telephone, so that the user can write what he wants to display on the touch screen while communicating with the voice, thus expressing the information of the caller more intuitively. The system is based on an embedded system and Ethernet is the data transmission medium, the system collects, encodes, transmits, receives, and decodes voice and graphic signals [29].

The working principle of the network multimedia communication system: the caller initiates a call, and after the two parties establish a connection, they can carry out a normal voice call, when they need to carry out a graphic communication, they can carry out a graphic communication through the touch screen. First of all, the sender inputs graphic information on the touch screen, and the graphic signal processing module collects the graphic signal and sends it to the microprocessor for processing, the microprocessor firstly judges whether the graphic data meets the standard of sending, and if it meets the standard, it compresses and encodes the data of one frame of the image, and then sends it out.

Design of Graphic Processing Module in Multimedia Communication

Graphic communication is one of the main functions of network multimedia systems. Its basic working principle is: first of all, the two communicating parties establish a dialogue, then the sender can write on the touch screen, the graphic processing module obtains the graphic data by communicating with the underlying touch screen driver and stores the graphic data in the format of a bitmap file in the buffer memory dynamically allocated by the program, then compresses and checks the graphic data and packs them up according to the network protocols in order, finally the processed digital graphic signal is transmitted to the network. Finally, the processed digital graphic signal is transmitted to the network. The receiver of the session receives the graphic data packets by listening to the port of the sender, then unpacks them in the opposite order of packing, decompresses the unpacked data by the compression protocol of the sender, and finally sends the digital signals to the liquid crystal display for display [30].

The graphic processing module can also be divided into two parts: graphic signal sender and receiver. The work to be done by the sending end of the graphic signal is mainly as follows: establishing the session; initializing and storing the bitmap; compressing and checking the graphic data; and sending the graphic data. The work to be done by the receiving end of the graphic signal includes establishing a session; initializing the bitmap; receiving and decompressing the graphic data; and displaying the graphic.

Deep Residual Adaptive Network Models

Image style transfer is the use of computer technology to convert natural images into specific style images. To quickly obtain transfer images with artistic style features, this paper adopts a deep convolutional residual neural network model with deep hierarchical construction based on the principle of deep learning. The abstract elements of image content and style are separated and extracted, and then these high-level abstract feature representations are processed using a pre-trained VGG deep convolutional model; an artistic effect image with original content and new style texture was synthesized through iterative optimization, effectively achieving the artistic effect of image style transfer.

Structure of Deep Residual Adaptive Network Model

This article proposes an image-style transfer method based on convolutional neural network texture synthesis. It is based on the Visual Geometry Group (VGG) network model and its random gradient descent mechanism. The loss function of the model network is designed, and after multiple iterations, the transfer image with artistic style is obtained.

Deep residual network: With the improvement of technical requirements, the model structure of deep learning networks gradually deepens, and the feature information extracted by neural network models becomes more and more abundant. As a result, neural networks are prone to problems such as vanishing or exploding gradients while the model structure deepens, resulting in a decrease in satisfaction with graphic style transfer. To address the shortcomings of the existing structure and alleviate the problems caused by gradient explosion and vanishing in the model, the deep residual network has added residual blocks [31] based on traditional network models. The residual block structure is shown in Figure 2.

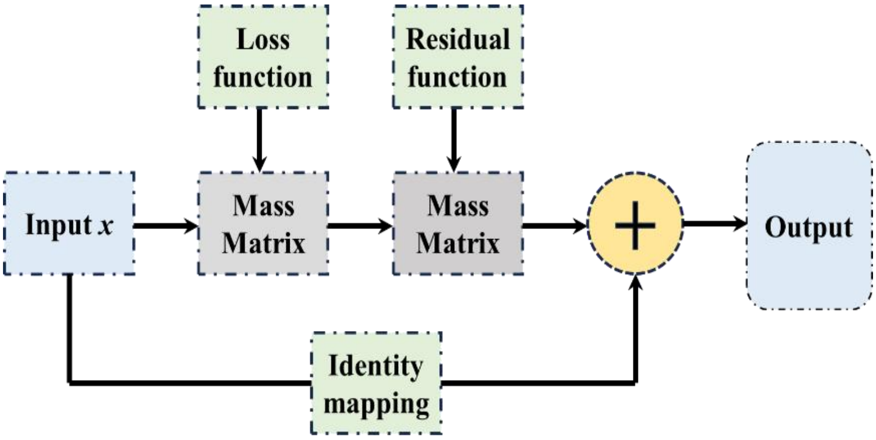

Figure 2. Residual Block Structure of Depth Residual Adaptive Network Model

Basic network model: A convolutional neural network is a classical neural network structure, which is mainly used for the processing and analysis of two-dimensional data such as images and videos. It can effectively extract the characteristic information of data, and it has laid the foundation for subsequent research in the field of deep learning [32], [33]. Its structural model mainly includes a convolution layer, pooling layer, activation layer, and full connection layer, which has excellent performance in image transfer learning [34], [35].

Convolutional layer: The convolution layer is the core layer of the convolution neural network. Convolution operation will be carried out in the convolution layer network to reduce the dimension of high-dimensional input data and extract the key features of data, so it is also called the feature extraction layer. In this network layer, neurons only need to connect with some neurons in the upper layer to reduce the dimension of data. This method is also called local connection.

Image convolution is generally in the form of discrete convolution, which is a linear transformation operation [36]. Discrete convolution has the characteristics of sparsity and parameter reuse; that is, in the convolution operation, the same set of parameters can be reused on different small blocks of the input image. This method can significantly reduce the number of parameters, thus reducing the computational complexity and storage requirements of the model. At the same time, sparsity can also improve the computational efficiency of the convolution layer. The discrete convolution operation formula is shown in Formula 1, and the convolution kernel operation process is shown in Figure 3.

$$y(n) = \sum_{i=-\infty}^{\infty} x(i)h(n-i) = x(n) \otimes h(n) \tag{1}$$

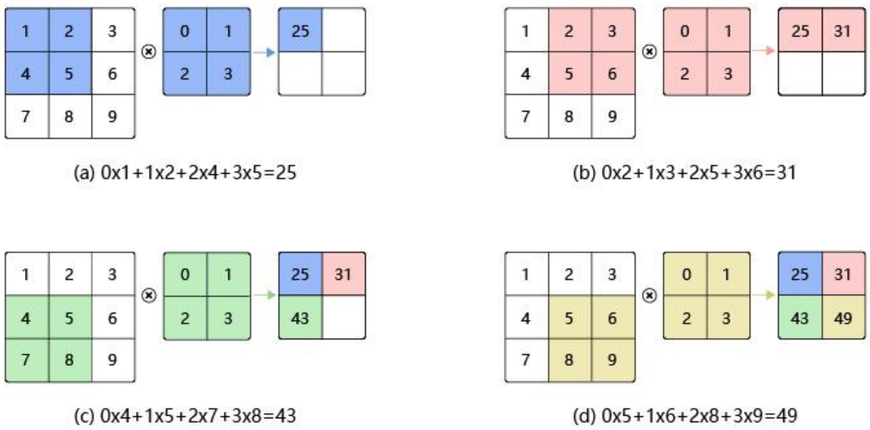

Figure 3. Schematic Diagram of Convolution Kernel Operation

Pool layer: The function of the pooling layer is to downsample the feature map generated after the above convolution operation. Currently, the widely used pooling methods mainly include maximum pooling [37] shown in Figure 4, and average pooling [38] shown in Figure 5.

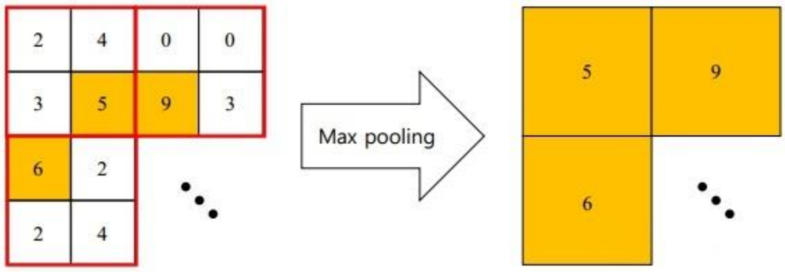

Figure 4. Schematic Diagram of Maximum Pooling

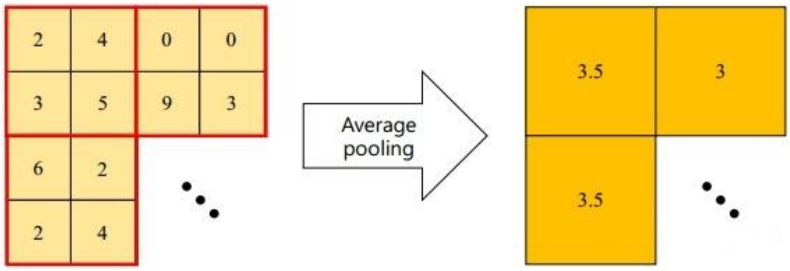

Figure 5. Schematic Diagram of Average Pooling

Because of the nonlinear characteristics of maximum pooling, it can better extract the salient features of the image and retain the edge information of the image. However, mean pooling simply calculates the average value of the image region, which may lead to the loss and blurring of information. Therefore, in the process of image transfer training in this paper, the maximum pool is selected as the pool layer in image transfer learning.

Active layer: To avoid the problem of gradient disappearance or gradient explosion in the neural network, we need to normalize the data in batches. Therefore, we need to select the activation function for the neural network. The common activation functions include sigmoid, tanh, and relu as shown in Figure 6.

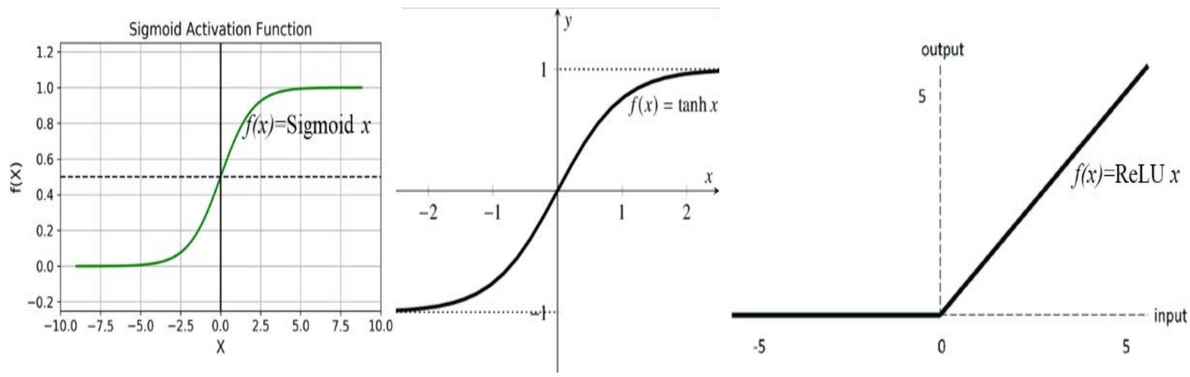

Figure 6. Common Activation Functions (Sigmoid, Tanh, TeLU)

Because when the activation function is selected in image migration, the real function is easier to converge in the training process and can better retain image features. The sigmoid function is easy to cause the gradient to disappear. Therefore, in image migration, it may be better to select relu as the activation function.

Full connection layer: Because the entire connection layer can map the learned "distributed feature representation" to the sample label space in the process of neural network model training and convert the features learned by the convolutional neural network into the results that can be classified, at the same time, in the process of fine-tuning, the whole connection layer can maintain a large model capacity to ensure the migration of model representation ability. Therefore, in the image transfer learning in this paper, to maintain a large model capacity and ensure the transfer of model representation ability, several total connection layers are added at the end of the model. The structure diagram of the primary network model in this paper is shown in Figure 7.

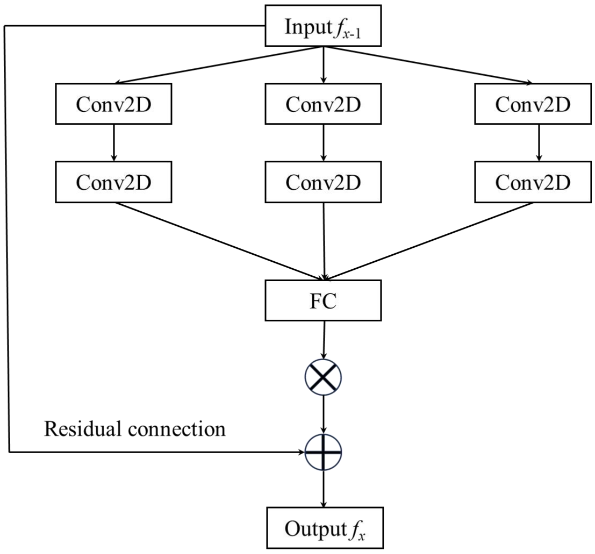

Figure 7. Schematic Diagram of Basic Network Model Structure

Establishment of Loss Function

Transfer learning is a machine learning method. Its core idea is to apply the trained model of one task to another task. To measure the difference between the composite image and the target image and guide the synthesis process, the loss function can be better defined by comparing the difference between the composite image and the target image in content, style, and other features. By minimizing the loss function, the composite image is as close to the target image as possible to achieve the effect of image migration, as shown in Figure 8.

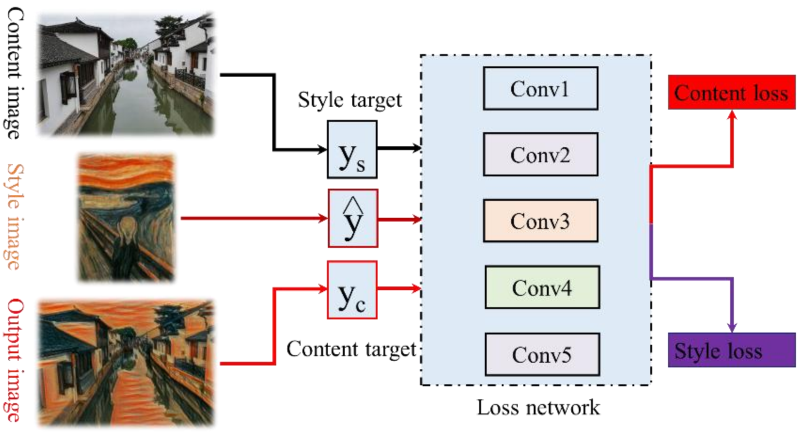

Figure 8. Schematic Diagram of Graphic Migration Effect Process

In style migration, the loss function is usually composed of the weighted sum of content loss, style loss, and total variation loss. The content loss measures the difference in content features between the composite image and the content image, the style loss measures the difference in style features between the composite image and the style image, and the total variation loss measures the smoothness of the composite image. By adjusting these weight super parameters, we can balance the relative importance of synthetic images in preserving content, migrating styles, and noise reduction.

Therefore, to quantify the difference between the composite image and the target image and guide the image migration process by minimizing the loss function, the following loss function needs to be constructed.

Step 1: define the input image as x and the output image as y. The convolution feature of the input image at layer l after the calculation is P^l_{ij}, where i represents the ith channel in the convolution layer and j represents the jth position in the convolution layer. Define the convolution feature of the output image y at layer l of the convolution layer as F^l_{ij}. After completing the definition of relevant variables, the content loss function is established as shown in formula 2:

$$Loss_{content}(x, y, l) = \frac{1}{2}\sum_{ij}(F^l_{ij} - P^l_{ij})^2 \quad (2)$$

The image style is often represented by the characteristic matrix Gram matrix of the image convolution layer, which is a symmetric matrix of the inner product of several groups of vectors. Consistent with the above, let the convolution feature be F^l_{ij}, and record the elements in the Gram matrix corresponding to this output as formula 3:

$$G^l_{ij} = \sum_k F^l_{ik} F^l_{jk} \tag{3}$$

Step 2: Let the height and width of the first layer of the convolution layer be ml and NL respectively, let the input image be x, the output image after style migration be y, the Gram matrix of the first layer of the input image x be X^l, and the Gram matrix of the first layer of the output image be Y^l. As in the previous section, a single-layer style loss function is established as shown in formula 4:

$$Loss_{style}(x, y, l) = \frac{1}{4N_l^2 M_l^2} \sum (X^l_{ij} - Y^l_{ij})^2 \tag{4}$$

Step 3: The total loss function of style in the loss network is shown in formula 5:

$$Loss_{style}(x, y) = \sum_l \omega_l \, Loss_{style}(x, y, l) \tag{5}$$

Thus, the total loss function can be established, as shown in formula 6:

$$Loss_{Total}(O_{style}, O_{content}, R) = \alpha Loss_{content}(O_{content}, R) + \beta Loss_{style}(O_{style}, R) \tag{6}$$

Where: O_{style} stands for the desired style image, $O_{content}$ represents the desired content image, and R represents the image to be generated; $Loss_{content}(O_{content}, R)$ stands for the content loss function, $Loss_{style}(O_{style}, R)$ stands for the style loss function; α, β are the super parameters to measure the proportion of the two loss functions in the total loss.

METHODOLOGY

Experimental Data Set

In the experimental process of this article, to make the experimental results representative, when selecting the dataset, the author randomly used the crawler software Python to crawl 1200 high-quality image sets with ultra-high resolution online as the training set of the model in this article. The crawled images mainly came from open-source images on Baidu Baike and Wikipedia, and each image had a resolution of nearly 2K. The images included image paintings, buildings, and portraits There are various categories such as landscape photos.

During the experiment, the image data was considered to be too large. Therefore, during the experiment, images with too large data in the original dataset were cropped to a certain ratio to reduce the burden of model training. At the same time, the cropped images were re-cleaned, mainly including image format conversion and geometric correction, to form a new dataset. At the same time, to more accurately describe the accuracy of the model in this article, in addition to using randomly crawled datasets for training, this article also selected the basic dataset ImageNet dataset for comparison and also used commonly used deep learning models VGG-19 and SRCNN for comparative testing. The above basic dataset not only contains rich image content but has also been widely validated by deep learning testers. The above basic model is also highly representative and has been fully demonstrated in the application of image style transfer. By comparison, the performance of this model can be comprehensively and effectively verified.

Experimental Parameter Setting

The experimental running environment of this paper is the TensorFlow deep learning framework, the experimental operating system is windows10 operating system, and the language environment is Python language. The experimental computer uses an 8-core 16-thread processor, more than 32GB memory, more than 512 GB SSD solid state drive and more than 2TB HDD mechanical hard disk, professional graphics card, more than 1200W power supply, and is configured with a rtx2080tigpu to support model operation.

Experimental Index Setting

To improve the generalization ability and robustness of the model in this paper, the training data are amplified by horizontal flipping, random rotation, and scaling. The training part is based on the Adam optimizer, and the initial value of the learning rate is set to 0.001. With the iterative training, the learning rate is adjusted and reduced by half every 1000 cycles. The total number of iterations is 20000, and the batch size of each iteration is 16. Due to the widespread use of PSNR in evaluating image quality and being a commonly used objective measurement method in this field, we chose PSNR as the first evaluation indicator for indicator evaluation. However, many experimental results have shown that the score of PSNR cannot be completely consistent with the visual quality seen by the human eye, mainly because those with higher PSNR appear to be worse than those with lower PSNR. Structural similarity can precisely compensate for this disadvantage and is also a commonly used indicator to measure image similarity. It can also be used to judge the quality of compressed images. Therefore, this article evaluates the model by combining two evaluation indicators: structural similarity (SSIM) and peak signal-to-noise ratio (PSNR).

Specific Implementation of the Experiment

This experiment is divided into two stages: the training stage and the generation stage. The experimental data used in each stage are different, as shown below.

Training phase: The data set used in the training phase includes two parts: style image and content image. The details are as follows. As shown in Figure 9, different pictures from various data sets are content pictures, which refer to the target processing picture. Relying on the content picture as the background, the style of this picture is changed to meet the needs of the creator; The function of a style picture is to assign the element content of the style picture to the content picture during training, to realize the graphic style transfer of content picture; Output picture refers to the target picture after the style transfer from the style picture to the content picture.

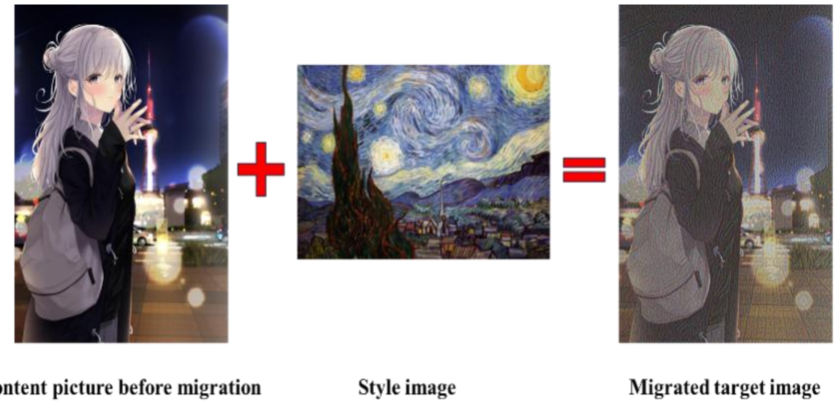

Figure 9. Schematic Diagram of Three Kinds of Pictures for Model Training

Generation stage: The trained model has been obtained in the generation stage. At this time, only the content image is needed. We choose natural landscape photos of different sizes as the content image, which will be embodied in the subsequent experiments.

RESULTS AND DISCUSSION

Multimedia Network Communication System Performance Testing

To test the system, two identical terminals are connected to the LAN through a router, with the system on the left as the sender and the one on the right as the receiver, with the address of the sender set to 192.168.1.6 and the address of the receiver set to 192.168.1.8. The results of the experiment are shown in Figure 10 below.

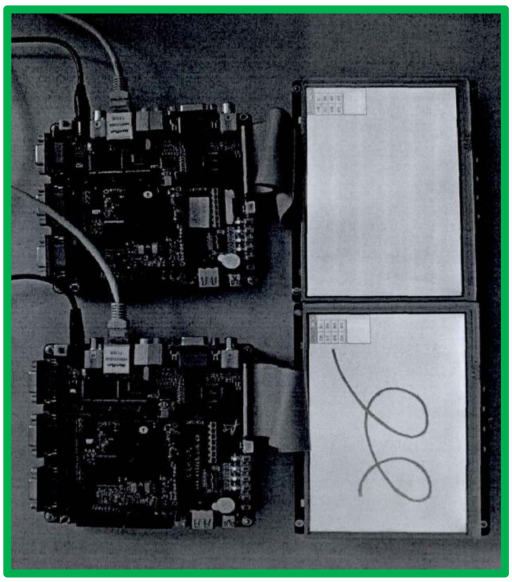

Figure 10. Experimental Effect

Deep Residual Adaptive Network Model Testing

The data set (hereinafter referred to as data set 1) and the basic data set ImageNet data set (hereinafter referred to as ImageNet data set) constructed in this paper are substituted into the depth residual adaptive network model (hereinafter referred to as model 1), the depth learning model vgg-19 (hereinafter referred to as egg-19) and the scan model (hereinafter referred to as scan) constructed in this paper for training, and two test results are obtained in Figure 11, Figure 12 and Tables 1 and 2, Figure 11 shows the image style migration results of the model built in this paper for different dataset images, and Figure 12 shows the image style migration results of other models for different dataset images. Table 1 corresponds to the PSNR index, and Table 2 corresponds to the SSIM index. The corresponding change charts are shown in Figures 11 and Figure 12.

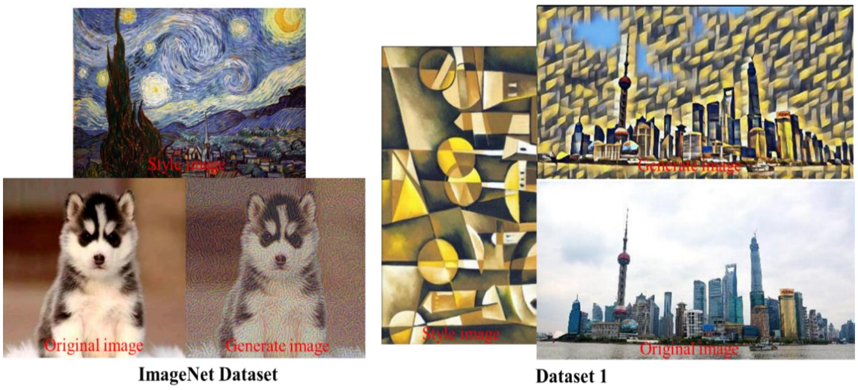

Figure 11. Image Style Migration Results of Model 1 for Images from Different Datasets

Based on Figure 11, it can be found that the ImageNet dataset is animal pictures, and Dataset 1 is Shanghai Bund exterior pictures. The image styles of the two are different, and the selected style pictures are slightly different. One is image painting, and the other is painting. However, from the performance of the training results of the two models, whether the ImageNet dataset is animal pictures or dataset 1 is Shanghai Bund exterior pictures, Model 1 can carry out a good style transfer on the photos of the above dataset, and apply the style of the style picture I want to the target picture, to achieve the purpose of style transfer, and the results are consistent. Refer to the detailed data in Table 1 for the specific results. It can be seen that this model has good application ability for the newly constructed dataset, which proves the feasibility of this model in the practice of image style transfer.

Figure 12. Image Style Migration Results of Other Models for Different Dataset Images

Based on Figure 12, it can be found that the training results of the two different training sets from the model VGG-19 and the model src are consistent with the model constructed in this paper. Whether the Imagenet dataset is animal pictures or dataset 1 is Shanghai Bund exterior pictures, the above comparison model can also carry out a good style migration for the above dataset photos, The style of the style picture I want is applied to the target picture, to achieve the purpose of style transfer. However, compared with model I, the VGG-19 model has more severe migration for animal models, preferring the style picture, while the src model for animal models is more inclined to the animal itself. Compared with model i in Figure 12 model i makes up for the shortcomings of both. The style migration is moderate, and can better express the author's ideas, The reliability of this model is confirmed.

Table 1. Comparison of PSNR Indexes of Different Models for Data Sets

Model	Dataset 1	Image Net Dataset
Model 1	36.35	36.09
VGG-19	31.33	31.24
SRCNN	32.48	32.01

Table 2. Comparison of SSIM Indicators of Different Models for Data Sets

Model	Dataset 1	Image Net Dataset
Model 1	0.9413	0.9207
VGG-19	0.9109	0.9287
SRCNN	0.8354	0.8698

According to the comparison results of the two evaluation indexes in Table 1 and Table 2, compared with the two traditional models vgg-19 and scan selected in this paper, the depth residual adaptive network model constructed in this paper is better than the former two models in terms of style transfer of graphics during the training process. According to the data in Table 1, in terms of peak signal-to-noise ratio, the model 1psnr constructed in this paper is 5.02db and 3.87db higher than other models for dataset 1, and 4.85db and 4.08db higher for Imagenet dataset, respectively; Through the data in Table 2, it is found that in terms of structural similarity, the model 1psnr constructed in this paper has increased by 0.0304 and 0.1059 for dataset 1, and -0.008 and 0.0509 for ImageNet dataset, respectively. From the data in Tables 1 and 2, it can be seen that the model constructed in this paper has strong graphic style migration ability for both dataset 1 and the ImageNet dataset dataset. Although the structural similarity of the ImageNet dataset is slightly lower than that of egg-19, it generally exceeds 0.9, showing a good graphic style migration level. The reason is that egg-19 has been used by many scholars to train the model, Therefore, its performance on the ImageNet dataset is better than that of the model constructed in this paper, but it is inferior to that of the model constructed in this paper on dataset 1. Therefore, it is not difficult to judge that the overall graphic style migration ability of model 1 constructed in this paper is stronger than that of other models and has a high migration level.

CONCLUSION

With the rapid development of wireless network technology and the rapid popularity of portable smart terminals, multimedia communication based on images and videos has become the favorite way of communication in the new era. Image style transfer technology is one of the research directions that has attracted much attention in the field of multimedia communication. To achieve the diversification and ease of use of images in the multimedia communication process, this paper researches the multimedia network communication technology and the image style transfer technology, aiming to achieve the mutual migration of different graphic styles through this technology and realize the simple technological artistic creation. The model introduces a convolutional neural network based on the unique characteristics of images, improves the convolutional neural network model by introducing an adaptive model and deep residual theory, and employs a combination of pixel loss and perceptual loss in the loss function to enhance the visual quality of the image reconstruction. We use the Python technique to grab different kinds of high-definition images to create a new dataset1 and introduce the traditional dataset ImageNet dataset and other VGG-19, SRCNN models to continue the comparison training. The results show that the reconstructed images are improved in subjective and objective evaluation indexes such as details and visual effects, and outperform the other two models in objective evaluation indexes such as PSNR and SSIM. At the same time, compared with the ImageNet dataset, the training effect of the model in this paper is comparable to that of the self-constructed dataset 1, which proves that the model constructed in this paper has good applicability, and that the model in this paper has strong practical value. The applicability of this technique in other fields has not yet been verified due to the limited depth of research and breadth of learning in other fields by the authors of this paper on related techniques. The authors of this paper will improve and optimize it in subsequent studies.

ETHICAL DECLARATION

Conflict of interest: No declaration required. **Financing:** No reporting required. **Peer review:** Double anonymous peer review.

REFERENCES

[1] S. Bera, S. N. Sur, and R. Bera, “Multimedia communication using DVB technology over open range,” *Procedia Computer Science*, vol. 70, pp. 282-288, 2015.

[2] E. Berrocal, I. V. Meglinski, D. A. Greenhalgh, and M. A. Linne, “Image transfer through the complex scattering turbid media,” *Laser Physics Letters*, vol. 3, no. 9, pp. 464-467, 2006.

[3] C. J. Wong, H. S. Lim, M. Z. MatJafri, and K. Abdullah, “Image data transfer over the Internet protocol for air quality studies,” in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII*, May 2006, vol. 6233, pp. 335-348.

[4] M. S. Slobodyan, “Methods of creation of permanent zirconium alloy joints in reactor art: A review,” *Tsvetnye Metally*, no. 10, pp. 91-98, 2016.

[5] B. Sunarto, “Basic knowledge and reasoning process in the art creation,” *Open Journal of Philosophy*, vol. 5, no. 5, pp. 285-296, 2015.

[6] T. Kinoshita, “A study in teaching art on the principle of creation: Searching for a method of teaching art,” *Art Education*, pp. 73-84, 2002, doi: 10.24455/aaej.23.0_73.

[7] C. K. Yang and H. L. Yang, “Realization of Seurat’s pointillism via non-photorealistic rendering,” *The Visual Computer*, vol. 24, pp. 303-322, 2008.

[8] A. D. Cheok, Z. S. Lim, and R. T. K. Tan, “Humanistic oriental art created using automated computer processing and non-photorealistic rendering,” *Computers & Graphics*, vol. 31, no. 2, pp. 280-291, 2007.

[9] S. J. Shin et al., “Style transfer strategy for developing a generalizable deep learning application in digital pathology,” *Computer Methods and Programs in Biomedicine*, vol. 198, 2021, doi: 10.1016/j.cmpb.2020.105815.

[10] M. Izadyyazdanabadi et al., “Fluorescence image histology pattern transformation using image style transfer,” *Frontiers in Oncology*, vol. 9, p. 519, 2019.

[11] M. Kumari, A. Khare, and P. Khare, “JPEG compression steganography & crypography using image-adaptation technique,” *Journal of Advances in Information Technology*, vol. 1, no. 3, pp. 141-145, 2010.

[12] Y. Li, Y. Wang, H. W. Tseng, H. Huang, and C .C. Chen, “Optimized and improved methods of image style transfer for local reinforcement,” *Sensors and Materials*, vol. 33, no. 9, pp. 3325-3332, 2021.

[13] F. Kolland, “National cultures and technology transfer: The influence of the Mexican life style on technology adaptation,” *International Journal of Intercultural Relations*, vol. 14, no. 3, pp. 319-336, 1990.

[14] X. Y. Qian, “Survey on non-photorealistic rendering technology for artistic images,” (in Chinese), *Journal of Engineering Graphics*, no. 1, p. 7, 2010.

[15] Y. Y. Liu, “A review of research on non-realistic rendering of computer images,” (in Chinese), *Information Communication*, no. 5, p. 2, 2013, doi: CNKI: SUN: HBYD.0.2013-05-023.

[16] Y. Zhai, “A review of research on non realistic rendering of computer images,” (in Chinese), *Journal of Jilin Radio and Television University*, no. 4, p. 2, 2015, doi: 10.3969/j.issn.1008-7508.2015.04.005.

[17] J. E. Kyprianidis, J. Collomosse, T. Wang, and T. Isenberg, “State of the ‘art’: A taxonomy of artistic stylization techniques for images and video,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 5, pp. 866-885, 2012.

[18] X. Cai, Y. Ge, R. Cai, and T. Guo, “Image color style migration method for mobile applications based color feature extraction,” *Journal of Computational Methods in Sciences and Engineering*, vol. 19, no. 4, pp. 879-890, 2019.

[19] J. Zhou, Y. Wang, J. Gong, G. Dong, and W. Ma, “Research on Image Style Convolution Neural Network Migration Based on Deep Hybrid Generation Model,” *Academic Journal of Computing & Information Science*, vol. 4, no. 8, pp. 83-89, 2021.

[20] L. A. Gatys, A. S. Ecker, and M. Bethge, “A neural algorithm of artistic style,” 2015, doi:10.48550/arXiv.1508.06576.

[21] L. A. Gatys, A. S. Ecker, and M. Bethge, “Image style transfer using convolutional neural networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2414-2423.

[22] Z. Qawaqneh, A. A. Mallouh, and B. D. Barkana, “Deep convolutional neural network for age estimation based on VGG-Face Model,” 2017, doi: 10.48550/arXiv.1709.01664.

[23] B. Talafha and M. Al-Ayyoub, “Just at VQA-Med: A VGG-Seq2Seq Model,” in *Conference and Labs of the Evaluation Forum*, Sep. 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:51940261

[24] S. K. Gorti and J. Ma, “Text-to-image-to-text translation using cycle consistent adversarial networks,” 2018, doi: 10.48550/arXiv.1808.04538.

[25] Y. Zhao, R. Wu, and H. Dong, “Unpaired image-to-image translation using adversarial consistency loss,” in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16*, Nov. 2020, pp. 800-815.

[26] M. Kaur, S. Satapathy, R. Soundrapandiyan, and J. Singh, “Targeted style transfer using cycle consistent generative adversarial networks with quantitative analysis of different loss functions,” *International Journal of Knowledge-*

Based and Intelligent Engineering Systems, vol. 22, no. 4, pp. 239-247, 2018.

[27]W. J. Li, S. J. Ruan, and D. S. Yang, “Implementation of energy — Efficient fast convolution algorithm for deep convolutional neural networks based on FPGA,” *Electronics Letters*, vol. 56, no. 10, pp. 485-488, 2020.

[28]G. Balachandran and J. Venu Gopala Krishnan, “Moving scene-based video segmentation using fast convolutional neural network integration of VGG-16 net deep learning architecture,” *International Journal of Modeling, Simulation, and Scientific Computing*, vol. 14, no. 1, p. 2341014, 2023.

[29]I. Ding Jr, N. W. Zheng, and M. C. Hsieh, “Hand gesture intention-based identity recognition using various recognition strategies incorporated with VGG convolution neural network-extracted deep learning features,” *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 4, pp. 7775-7788, 2021.

[30]Y. Tai, J. Yang, and X. Liu, “Image super-resolution via deep recursive residual network,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3147-3155.

[31]K. Hendling, G. Franzl, B. Statovci-Halimi, and A. Halimi, “Residual network and link capacity weighting for efficient traffic engineering in MPLS networks,” *Teletraffic Science and Engineering*, vol. 5, pp. 51-60, 2003.

[32]P. Kim, (2017). “Convolutional neural network,” in *MATLAB Deep Learning*, Berkeley, CA: Apress, 2017, pp. 121-174, doi: 10.1007/978-1-4842-2845-6_6.

[33]Y. Chen, Z. Cao, J. Zhang, Y. Liu, D. Yu, and X. Guo, “Wearable ultraviolet sensor based on convolutional neural network image processing method,” *Sensors and Actuators A: Physical*, vol. 338, p. 113402, 2022.

[34]Y. Chen, “Research on convolutional neural network image recognition algorithm based on computer big data,” *Journal of Physics: Conference Series*, vol. 1744, no. 2, p. 022096, 2021.

[35]Z. M. Zain, S. Sakri, N. H. Asmak Ismail, and R. M. Parizi, “Software defect prediction harnessing on multi 1-dimensional convolutional neural network structure,” *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1521-1546, 2022, doi: 10.32604/cmc.2022.022085.

[36]J. Kim, J. K. Kang, and Y. Kim, “A research on quantized lightweight convolutional neural network structure for facial emotion recognition,” *Journal of the Institute of Electronics and Information Engineers*, vol. 57, no. 12, pp. 51-59, 2020.

[37]M. Tan, F. Wu, D. Kong, and X. Mao, “Automatic liver segmentation using 3D convolutional neural networks with a hybrid loss function,” *Medical Physics*, vol. 48, no. 4, pp. 1707-1719, 2021.

[38]L. A. Kumar and M. E. Jebarani, “An optimized convolutional neural network for multi-spectral change detection technique,” *Journal of Uncertain Systems*, vol. 15, no. 3, p. 2241006, 2022.