Creation mechanism of new media art combining artificial intelligence and internet of things technology in a metaverse environment

Xiaolong Wang[1] · Ling Cai[1] · Yunhao Xu[2]

Abstract

The Metaverse is regarded as a brand-new virtual society constructed by deep media, and the new media art produced by new media technology will gradually replace the traditional art form and play an important role in the infinite Metaverse in the future. The maturity of the new media art creation mechanism must also depend on the help of artificial intelligence (AI) and Internet of Things (IoT) technology. The purpose of this study is to explore the image style transfer of digital painting art in new media art, that is, to reshape the image style by neural network technology in AI based on retaining the semantic information of the original image. Based on neural style transfer, an image style conversion method based on feature synthesis is proposed. Using the feature mapping of content image and style image and combining the advantages of traditional texture synthesis, a richer multi-style target feature mapping is synthesized. Then, the inverse transformation of target feature mapping is restored to an image to realize style transformation. In addition, the research results are analyzed. Under the background of integrating AI and IoT, the creation mechanism of new media art is optimized. Regarding digital art style transformation, the Tensorflow program framework is used for simulation verification and performance evaluation. The experimental results show that the image style transfer method based on feature synthesis proposed in this study can make the image texture more reasonably distributed, and can change the style texture by retaining more semantic structure content of the original image, thus generating richer artistic effects, and having better interactivity and local controllability. It can provide theoretical help and reference for developing new media art creation mechanisms.

Keywords Metaverse · Artificial intelligence · Internet of Things · New media art · Image style transfer · Texture synthesis · Deep neural network

Extended author information available on the last page of the article

1 Introduction

At present, the new frontier of science and technology development is intelligence. Especially in recent years, the rapid development of artificial intelligence (AI) technology, big data acquisition, and Internet of Things (IoT) technology has produced various intelligent products and devices for people, bringing many conveniences and changes to life. Moreover, in 2021, a brand-new concept of the Metaverse was developed. The Metaverse is likely to be valued by everyone as another world of human beings in the future, and the construction of Metaverse also has a huge workload, which needs to meet people's needs in many aspects, such as vision and hearing.

With the development of science and technology, new media art can just meet the basic construction of the Metaverse, especially digital art [30]. It is also the inevitable development path and new creation form of the digital age to assist visual art creation in digital art through AI-related technologies [35]. Through this new media form and language, the expression of painting art can be greatly enriched, and a new visual form can be obtained [5].

New media art is an interdisciplinary creative field that uses digital technologies and interactive media to create and express artworks. This includes, but is not limited to, digital painting, virtual reality (VR), interactive art, web art, and computational art. A core feature of new media art is using digital tools and interactivity to create works that enable the audience to participate in and interact with the work, thus broadening the boundaries of art and ways of expression [8]. With the development of AI and IoT technology, the creative methods and experiential forms of new media art have also undergone profound changes. Through the application of AI-related technologies, artists can use big data and algorithms to achieve special effects such as image generation, style conversion, and content creation, providing more creative possibilities. Meanwhile, with the development of IoT technology, new media artworks can interact with the audience in real-time, and two-way communication and participation between the works and the audience can be realized through technical means such as sensors, sensors, and networks.

In the Metaverse environment, new media art has a broad application prospect. Metaverse, a virtual world based on digital technology, provides a more open and diverse creative platform for new media art. In this virtual space, artists can connect the real world and the virtual world through IoT technology, breaking time and space restrictions and creating more diverse artistic experiences. Meanwhile, through the application of AI technology, artworks can make real-time changes and adjustments according to the audience's needs and emotional feedback, making the interaction between art and audience closer and more personalized.

The main meaning and contribution of this study are to integrate AI and IoT technology in the Metaverse environment, optimize the creative mechanism of new media art, and realize the style transformation of content result images. The motivation is to utilize AI and IoT technology to further improve convolutional neural network (CNN) and visual texture synthesis technology, and finally

promote cutting-edge research in the field of new media art. The main research direction and research method are to generate a confrontation network and deep learning (DL) model, and to explore the application in image generation, style transformation, and content creation in the Metaverse environment. With the support of big data and algorithms, innovative technologies are developed to realize the automatic generation and adaptation of artistic works, thus providing artists with more creative possibilities. Secondly, with the help of IoT technology, a bridge connecting the real world and the virtual world is established. Through technical means such as sensors, sensors, and networks, people have established real-time interaction and participation between artworks and the audience. The audience's behavior and emotional feedback will become the basis for the adjustment and change of artistic works, and this personalized interaction will enrich and deepen the artistic experience.

The organizational structure of this study is as follows. Section 1 is an introduction that introduces the emergence of Metaverse and the technical background of IoT development, while also elaborating on the innovation and contribution of this study. Section 2 is the literature review, which summarizes the research literature on new media painting art. Section 3 is the method, which uses visual geometry and texture synthesis technology to design a style transfer method based on feature synthesis, and applies it to Metaverse environment. Section 4 is the results and discussion, which discusses the main results of the study by collecting experimental datasets and analyzing the experimental environment. Section 5 is the conclusion, which draws research conclusions through the analysis and induction of experimental results.

2 Literature review

2.1 Research status of computer vision applied to an image style transformation

Academic circles made many achievements in researching digital painting art and image style transfer. For example, Hien et al. thought that there were some limitations in the output quality and implementation time of neural style transfer, so they put forward an image transformation network to generate higher quality artworks, and have a higher ability to perform on a larger number of images [16]. Huang et al. aimed to use computer support to help people's visual perception in painting practice. The convolution neural network model was adopted to recognize what people draw. Then, the computer graphics were output according to the recognition results to support the creation of mosaic design patterns [17]. Nasri and Huang studied the defects of the area where the paint layer was lost in ancient murals. They proposed an effective method to automatically extract defects and map degradation from the Red Green Blue (RGB) images of ancient murals in Bey Palace, Algeria. Firstly, the dark channel prior was used for preprocessing to improve the quality and visibility of murals. Secondly, by calculating the average and variance of all classes in three frequency bands and the covariance matrix of all classes, the training samples and pixel grouping were assigned to the nearest samples based on Mahalanobis distance

to realize the cavity extraction of murals. Finally, the accuracy of extraction was calculated [26]. Castellano and Vessio outlined some of the most relevant DL methods for pattern extraction and recognition in visual arts. The research results showed that the increasing number of DL, computer vision, and large digital visual arts collections provided new opportunities for computer science researchers, helped the art world to use automatic tools to analyze and further understand visual arts, and also supported the spread of culture [6]. Ahmadkhani and Moghaddam examined the influence of using image style on the performance of social image recommendation systems. First, an image style recognition method based on depth feature and compact convolution transform was proposed. Second, an image recommendation system based on image style recognition was introduced. The experimental results denoted that style had a positive effect, and the percentage of personalized image recommendations can be improved by 10% [1].

2.2 Research status of the integration of AI and IoT technology in the field of new media art

The application of AI and IoT technology integration in the field of new media art reflects the current trend of mutual penetration of science and art, thus opening up a broader field of interaction for new media art. Many scholars have carried out related research. Radanliev et al. [29] studied AI and IoT technology in Industry 4.0, and put forward a new design of AI in the network physical system through qualitative and empirical analysis of the framework of Industry 4.0. Grounded theory was adopted to analyze and model the edge components and automation model in the network physical system, and a new comprehensive framework was designed for the future network physical system, which had practical application value for the deep integration of AI and IoT [29]. Radanliev et al. [28] explored the emerging data and technical forms, and optimized data collection and multimedia data forms through bibliometrics review. The research results revealed that new data storage methods were needed to avoid network technology risks brought about by VR technology [28]. Esenogho et al. [10] examined the structure of a new generation of smart grid integrating AI, IoT, and fifth-generation mobile communication, aiming at enhancing the interconnection of power generation, transmission, and distribution networks by introducing the paradigm of the traditional power grid. The research had practical reference value for enhancing the self-repair and perception function of smart grid [10]. Liao et al. [21] investigated the semantic context-aware image style conversion technology, through semantic context matching, and then the hierarchical local-to-global network architecture. Their exploration could solve the inconsistency between different corresponding semantic regions [21]. To sum up, the discussion on the optimization between AI technology and industrial IoT reflected the innovation of new media art creation mechanisms. This study investigated how integrating AI and IoT technology can promote the remodeling and transfer of image style in the Metaverse environment. AI mainly studied human intelligence activities, learning from its characteristics and rules, and using computer hardware and software to construct an artificial system with a certain degree of intelligence to accomplish more

tasks that require human intelligence to be competent [9]. In recent decades, great progress has been made in many fields, such as natural language processing (NLP), image speech recognition, and so on, and they have all been put into practical application. The technical core of realizing AI is machine learning (ML) and DL [15].

3 Summary

To sum up, many scholars made remarkable progress in digital painting art and image style transfer. At the same time, some scholars proposed the integration of AI and industrial IoT to facilitate the reshaping and changing of image styles in the Metaverse environment. For example, Radanliev et al. [29] designed a comprehensive framework for future cyber-physical systems through qualitative and empirical analysis of the framework of Industry 4.0, which had practical application value for realizing the deep integration of AI and IoT. Thus, by comparing and analyzing the literature review, it was possible to discover the application and innovation of AI and IoT technologies in the field of new media art. However, Metaverse, which was currently built using AI and IoT, still cannot flexibly change its image style. Therefore, the main contribution of this study is to promote the change of image style in the Metaverse environment through the integration of AI and IoT technology, to optimize the creation of new media art. By improving CNN and visual texture synthesis technology, the image style transfer method based on feature synthesis can be successfully realized, and higher quality and diversified new media artworks can be provided to enrich the form and content of artistic expression. Meanwhile, the objectivity and reliability of the experimental results were verified through the comprehensive use of visual evaluation and subjective evaluation methods. The exploration and demonstration of this study provided a practical and innovative reference for the integration of new media art with AI and IoT technology, and also opened up a new direction for the development and exploration of new media art in the future.

4 Research methodology

4.1 Artificial intelligence and Internet of Things

AI is a field of computer science that represents algorithms and systems capable of simulating human intelligent behavior, including techniques such as ML, DL, NLP, and computer vision [33]. ML is the software core of the basic technology of AI, which mainly studies computer simulation of some behaviors of human learning to achieve the purpose of absorbing new knowledge and skills [4]. In recent years, with the development of computer hardware technology, all kinds of powerful chips have been used in DL, thus greatly increasing the training speed [12]. The work of AI in digital art is displayed in Fig. 1.

Figure 1 is a painting, "Space Opera House" generated by AI in August 2022, which won the first prize in the digital art competition of the Colorado Expo [11]. It also shows that the current AI technology has made rich achievements in

Fig. 1 AI generates paintings: Théâtre D'opéra Spatial (image from the network)

information search, data mining, image processing, and other fields [7]. According to the constructed image style transfer model, this study applies the convolution neural network, which is common in DL.

IoT is a technology that connects the physical world with the digital world, enabling the collection, transmission, and analysis of real-time data through sensors, devices, and networks [20]. Its technical fields cover radio frequency identification, sensors, embedded software, and transmission data calculation, while its products and services cover smart homes, transportation and logistics, and personal health [22]. The main function of the IoT is intelligent perception, which can form a large-scale dynamic internet between things and people, thus saving the time and cost of traditional human resources [34]. As the most important medium of cultural communication, visual language plays a vital role in every IoT system. Consequently, the development of the IoT system and new media art must complement each other, and the future development of the metaverse vision field must also be inseparable from the intelligent perception of everything in the real world by the IoT [23].

4.2 Visual geometry group and texture synthesis technology

As a representative algorithm of DL, CNN belongs to feedforward neural networks, and its biggest feature is that it includes convolution calculation. The common CNN consists of the following three parts: input layer, hidden layer, and output layer. The hidden layer is a very critical part, which contains a convolutional layer, a pooling layer, and a fully connected (FC) layer [27]. In the input layer, one-dimensional

CNN can receive one-dimensional data of spectral sampling or time sampling and arrays of two-dimensional image pixels. In contrast, two-dimensional CNN can receive two-dimensional and three-dimensional arrays of image pixels or multiple channels [25]. Additionally, CNN has three structural characteristics, which are local connectivity, weight sharing, and space pooling. Among them, local connectivity means that each neuron is only connected to a local area of input data, and the spatial size of this connection is called the receptive field of neurons. Weight sharing means that parameter sharing controls the number of parameters in the convolutional layer. The function of spatial pooling is to gradually reduce the spatial size of data. The common CNN is plotted in Fig. 2.

Figure 2 denotes that there are usually many convolution kernels in each convolutional layer that contain matrix operations between elements, and each element corresponds to a weight coefficient and deviation term. The main function of the convolution kernel is to extract data features. During the operation of CNN, the convolution kernel will slide on the two-dimensional matrix and perform convolution calculations with the covered position [18]. The convolution neural network used in this study is the visual geometry group (VGG), which the Department of Science and Engineering of Oxford University researches. It is applied to generate and restore the characteristic graph. Its biggest feature is that the small convolution kernel replaces the common large convolution kernel. On the one hand, it can increase the nonlinear mapping. On the other hand, it can sparse the network parameters and reduce the running pressure [32].

Traditional texture synthesis algorithms are classified into global statistics-based and nonparametric stitching. Global statistics is employed to define statistical data in a random process to express the texture structure of samples. It is more suitable for generating disordered texture images, but it is not effective for repetitive texture
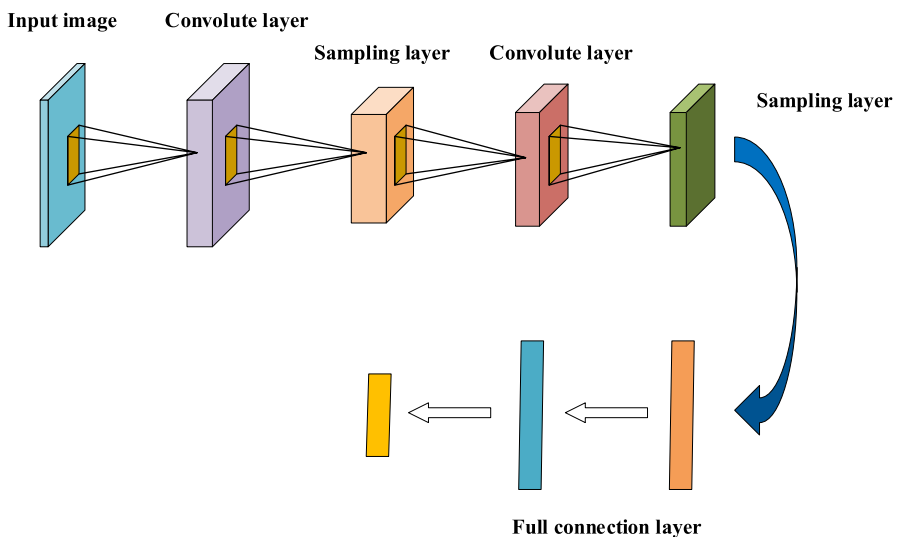

Fig. 2 Schematic diagram of CNN structure

or fine texture [19]. Nonparametric stitching methods are divided into spatial coherence texture synthesis method and block-based texture synthesis method. The feature synthesis method in this study is also based on the spatial coherence texture synthesis method. The specific synthesis process is shown in Fig. 3.

Figure 3 illustrates the sequential synthesis of output images following the order of scanning lines. In this context, the blue segment of the output image represents the synthesized portion. The next pixel in the synthesis sequence is determined based on the pixels within the nearby L-shaped neighborhood that has already been synthesized. A color error calculation is performed between the current point and each pixel in the L-shaped neighborhood found in the candidate pixel list from the input image. This calculation is used to identify the most suitable pixel, which is then copied to contribute to the image synthesis process. This approach offers notable advantages, including high execution efficiency and improved texture integrity [36].

In the texture construction of neural style transfer (NST), the initial step involves computing the differences among the target image, content image, and style image within the feature map. Subsequently, an optimization process aims to minimize the overall differences. The most commonly used method is to directly optimize and iterate the target image, taking the initial target image as the content image. A deep convolutional neural network is then utilized to extract diverse levels of depth image content and style feature information. Optimization of the target image is carried out through the Gram matrix. This process involves applying the gradient descent method for backpropagation to achieve an enhanced style conversion effect [31]. However, image reconstruction is accomplished through an inverse convolutional neural network, primarily focused on reversing the feature information within the convolution network. This is achieved by minimizing the loss in feature
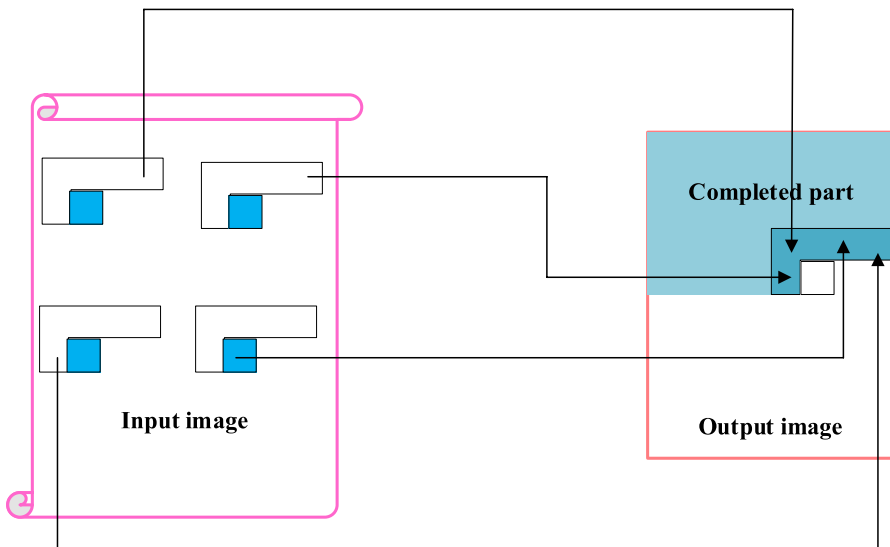

Fig. 3 Schematic diagram of spatial coherence texture synthesis process

reconstruction and training the given feature with the inverse convolutional neural network to obtain the training image. Subsequently, a weighted averaging operation is applied to obtain the pre-image, thus reinforcing the capacity for expressing image features [37].

4.3 Construction of image style transfer model based on feature synthesis in the Metaverse environment

In this study, a style transfer method based on feature synthesis is constructed, which mainly converts the image style transfer process into the constructed feature space. Through the two feature maps of content and style images, and taking into account the excellent features of traditional texture synthesis, the target image feature map is obtained. The flowchart based on feature synthesis is illustrated in Fig. 4.

In Fig. 4, first, the above VGG19 (very deep convolutional networks for large-scale image recognition) network architecture is used in the flowchart to obtain feature images of different levels of content and style. Second, the special diagnosis map of the target image is acquired by combining the two feature images through a synthesis scheme. Lastly, it is reversed into the target image. In the choice of model, the standard of model selection is the size of the hidden layer and the number of convolutional layers. The reason for choosing the VGG19 network is that the model structure is very simple, and the whole network uses the same convolution kernel size, so the model's performance can be improved by deepening the network structure. Among them, feature synthesis is a method that takes the content image as the constraint condition, and the style image provides the feature sampling points, and synthesizes the target feature map that contains both the texture information of the style image and the semantics of the content image [39]. Because the neural network is in the shallow layer and the middle layer, the image quality obtained by feature transformation is better. Therefore, combining the advantages of shallow and deep neural levels, a feature synthesis method from deep to shallow is constructed, as suggested in Fig. 5.

In Fig. 5, the constructed image style transfer model based on feature synthesis mainly includes five steps. The first step is to select the initial layer and the target layer. The appropriate DL neural network is chosen, and the initial layer and the
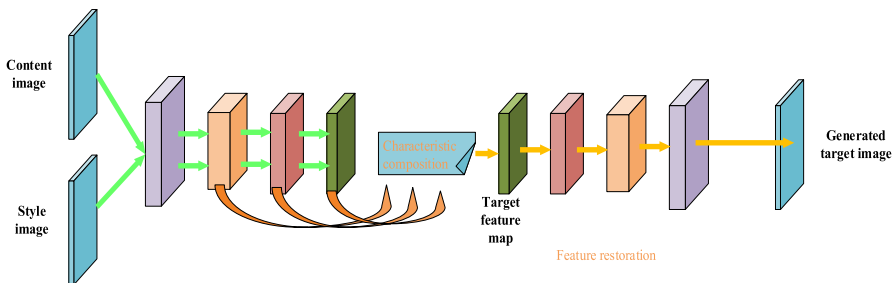

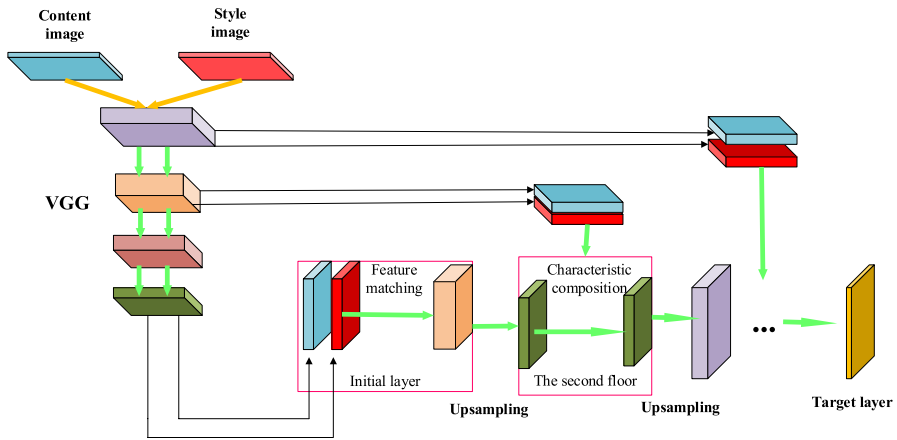**Fig. 4** Flowchart of feature-based synthesis

Fig. 5 Schematic diagram of layer-by-layer feature synthesis method from deep to shallow

target layer are selected from this network to obtain the content and style informa-
tion of the input image. The second step is to obtain the initial feature map. The
greedy search algorithm is used to adjust the pixel values on the initial layer to min-
imize the difference between the feature maps of the content image and the style
image, thus generating the feature maps of the initial layer. The third step is the
migration of feature points. The feature points in the style feature map are copied to
the corresponding coordinate positions in the initial feature map. The fourth step is
texture synthesis. The coherent texture synthesis method is employed to adjust the
pixel value of the image and generate the texture of the image by minimizing the
semantic difference between the style and content feature map and the initial feature
map. The fifth step is iterative synthesis. Through several iterations, the feature map
is synthesized step by step to better match the semantics of the content image. In
each iteration, steps 3 and 4 are repeated, constantly improving the resulting image
until a satisfactory result is obtained.

AI and IoT technologies are further integrated into the virtual environment of
Metaverse to provide personalized virtual assistants, and the constructed image
style transfer model is applied to meet users' needs and improve their virtual experi-
ence. At the same time, IoT sensors and devices can capture real-world data such as
sound, images, and motion. Combining AI's voice and image style transfer models
ultimately enables real-time environment perception and interaction, enabling users
to interact more naturally with virtual environments.

4.4 Datasets collection and experimental environment

This study mainly explores using different model structures to obtain target images
of style transfer from original images of different styles. It randomly selects 2000
images involving people, plants, buildings, animals, and landscapes from the
world's largest ImageNet dataset (https://image-net.org/) as the experimental data-
set. According to the ratio of 7:3, it is divided into the training set and the test set.

The experimental environment is a desktop computer equipped with the Ubuntu 20.04 system, GTX1070 graphics card, and 12 GB memory. Moreover, the Tensorflow experimental framework is used for training and the style network model of carbon painting and calligraphy is used. In terms of specific experimental design, 2000 images using the ImageNet database are first randomly selected, covering people, plants, buildings, animals, and landscapes. To increase the diversity of datasets, the Open Images database is introduced to provide more diverse image data, which can cover more types of images and styles and make the research results more comprehensive. In addition, the existing dataset can be enhanced by data amplification technology, and the transformation operations such as rotation, translation, scaling, and deformation can be applied to the image, or noise and blur can be introduced to increase the diversity and generalization ability of the data.

Next, the dataset is divided into training sets and test sets according to the ratio of 7:3 to ensure the generalization ability of the model. The experimental environment uses a desktop computer with Ubuntu 20.04 system, GTX1070 graphics card, and 12 GB memory. This configuration can provide enough computing resources and storage space to support the training and testing of the model. The experimental framework used in the experiment is Tensorflow, which is an open-source framework widely used in DL tasks. Tensorflow offers a wealth of tools and libraries, which can be used for model training, adjustment, and evaluation conveniently.

4.5 Parameters setting

To compare and analyze different types of style transfer images, this study trains a number of different style transfer network models. First, the target layer is fixed at the conv3-1 layer, the L-shaped neighborhood size is set to 5×5, and the content weight is set to 0.3. Then, the initial layer is changed to conv5-1, conv4-1, and conv3-1, respectively.

When the number of layers between the initial layer and the target layer increases, the network level will deepen, making it easier to get large irregular feature blocks, and the obtained image will have stronger style features. If the number of layers decreases, the style features will weaken, but the corresponding content features will be stronger. Therefore, it can be concluded that by changing the network depth of feature synthesis, the style expression effect of feature synthesis can be changed to some extent. Then, the weight parameter of content error is modified, which is mainly used to balance content loss error and style loss error. The target layer is conv3-1, and the initial layer is conv4-1.

As the weight coefficient increases, the image's content quality improves, while the texture feature block diminishes in size, resulting in a reduction in the effectiveness of the image's stylistic expression. On the contrary, a decrease in weight enhances the texture integrity of the target image and strengthens the capacity to convey style features. However, this comes at the expense of reduced content feature expression. Based on these parameters, achieving a more abstract style sample requires reducing the weight assigned to content error and increasing the depth of the initial feature layer. Conversely, if a more realistic style sample is desired,

the content error weight should be increased, and a shallower initial feature layer employed.

To further explore the performance and effect of different style conversion network models, the parameters of the models are also adjusted and optimized. The main adjusted parameters include the number of network layers, learning rate, iteration times, batch size, etc. Table 1 exhibits the different parameter combinations and corresponding model names used here:

Table 1 shows the M1–M6 model. The number of network layers, learning rate, iterations, and batch size of each model differ. In the M1–M6 model, a new image generation method is synthesized by using the constructed style transfer method based on feature synthesis. The specific model adjustment process and steps of the style transfer method of feature synthesis used here are as follows. First, a pre-trained CNN model is selected. This study uses VGG19, which is a deep network composed of 19 convolutional layers, FC layers, and pooled layers, and can effectively extract high-level semantic features and low-level texture features of images. Second, the feature representations of content images, style images, and generated images at different levels in the CNN model are determined. This study selects conv3-1 as the content target layer, and conv1-1, conv2-1, conv3-1, conv4-1, and conv5-1 as the style target layer. The content target layer is employed to retain the structural information of the content image, and the style target layer is used to capture the texture information of the style image. Subsequently, content loss and style loss functions are established to assess the distinctions in features between the generated image and the content and style images on the designated layer. In this study, the content loss function adopts the mean square error, while the style loss function employs the Gram matrix. The Gram matrix method quantifies feature correlations, providing insights into the statistical distribution characteristics of style images. Then, the total loss function is defined, which is the weighted sum of the content loss function and the style loss function. In this study, a content weight parameter α and a style weight parameter β are used to control the balance between content loss and style loss. Different values of α and β will lead to different style conversion effects. Finally, the backpropagation algorithm is adopted to optimize the generated image to minimize the total loss function. Here, the random gradient descent method is used as the optimizer, and a learning rate parameter γ is set to control the optimization speed. Different γ values will affect the stability and convergence of the optimization process.

Table 1 Model parameter adjustment values of style transfer method based on feature synthesis

Model name	Network layers	Learning rate	Iterations	Batch size
M1	5	0.001	1000	8
M2	4	0.003	1500	16
M3	3	0.001	2000	8
M4	5	0.002	1000	16
M5	5	0.001	2000	16
M6	5	0.002	1500	32

To evaluate the performance of the proposed model, the image style transfer accuracy of M1–M5 set at different parameters of the model is compared. Further, the constructed image style transfer model based on feature synthesis is compared with the style transfer image recognition accuracy of the model algorithm proposed by CNN, AlexNet, and related scholars Ahmadkhani & Moghaddam [1]. Furthermore, in the dimension of generating effect, the constructed model is compared with the automatic image generation effect using IoT.

5 Experimental Design and Performance Evaluation

5.1 Accuracy recognition analysis of image style transfer based on different model algorithms

The image style transfer accuracy of the proposed model for M1–M6 with different parameter settings is compared, as presented in Fig. 6. The constructed image style transfer model based on feature synthesis is compared with the style transfer image recognition accuracy of the model algorithm proposed by CNN, AlexNet, and related scholars Ahmadkhani and Moghaddam [1], as indicated in Fig. 7.

In Fig. 6, comparing the image style transfer accuracy of M1–M6 with different parameter settings, it can be found that the M5 model has the best image style transfer accuracy recognition, reaching 97.48%. Moreover, the recognition accuracy of each model after image style transfer is in descending order from M5 > M1 > M4 > M6 > M2 > M3. Among them, the M3 model has the lowest recognition accuracy of image style transfer, reaching 83.51%. From this, after the image style transfer of M1-M6 with different parameters set, the recognition accuracy is relatively high. This can offer more accurate support for the intelligent development of the new media art field.

In Fig. 7, the constructed image style transfer model based on feature synthesis is compared with the style transfer image recognition accuracy of the model algorithm

Fig. 6 Comparison results of image style transfer accuracy recognition for M1-M6
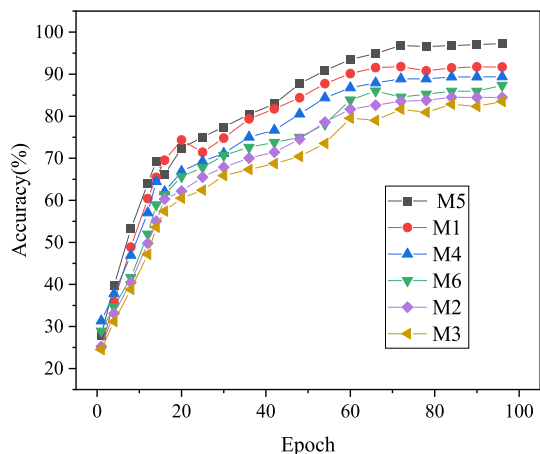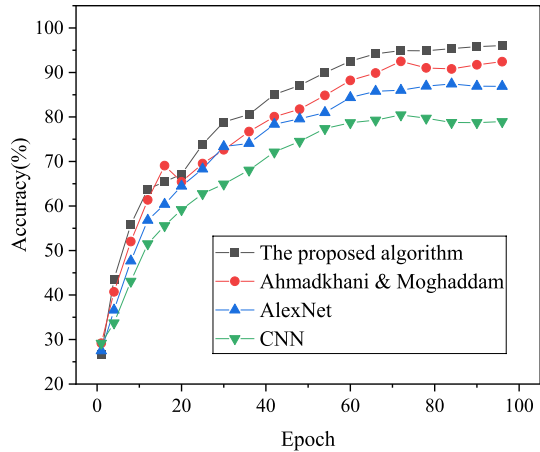
Fig. 7 Comparison results of image style transfer accuracy recognition under different model algorithms

proposed by CNN, AlexNet and scholars Ahmadkhani and Moghaddam [1]. It can be found that compared with other model algorithms, the proposed model algorithm is significantly higher, and the image recognition accuracy exceeds 96.03%, which is at least 3.61% higher than the algorithm adopted by other scholars (such as CNN, AlexNet, etc.). The recognition accuracy of each algorithm in order from the highest to the lowest is the proposed algorithm > the algorithm proposed by Ahmadkhani and Moghaddam [1] > AlexNet > CNN. Compared with the algorithms adopted by other scholars, the constructed model has a higher image recognition accuracy, Thus, the constructed model can better capture the information of images after style transfer, and provide more accurate support for the intelligent development of image interaction in the Metaverse environment.

5.2 Comparative analysis of image generation effect

In the process of performance evaluation of new media art images, it is necessary to list the objective evaluation criteria in detail first. Here, the network structure model of VGG19 is used to extract image features. Then, after comparing the brightness, contrast, and structure of images and synthesizing different structural similarity measurement schemes, the performance scores of similar images under different processing schemes are measured and compared by calculating the brightness and contrast of local patterns. Because the evaluation of image quality in the computer field is unsuitable for the aesthetic value of the image, this study analyzes the image style and image artistry to compare the effect of image style transmission in different research results. The style comparison and artistic comparison of the same image generated by different models are portrayed in Figs. 8 and 9.

Figure 8a and b depicts style transformations of the same image generated by distinct models. Figure 8a adopts an automatic image generation model based on the IoT, and the style of the generated image is abstract expressionism. Conversely, Fig. 8b utilizes an image style conversion model based on feature synthesis, yielding a generated image with an impressionistic style. Abstract expressionism, which
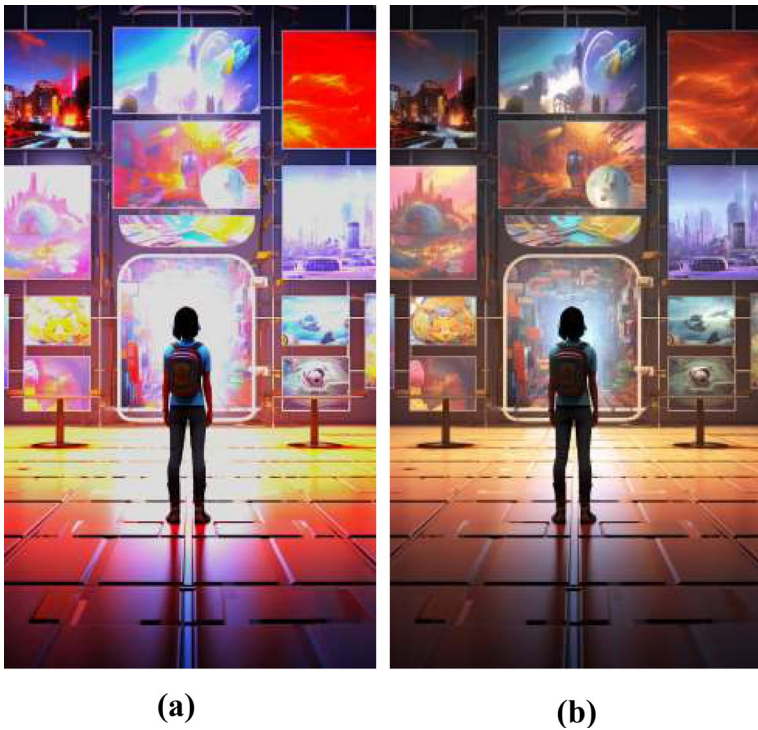
Fig. 8 Style display and comparison of artistic images of the same new media generated by distinct models (**a** automatic image generation model based on IoT; **b** image style conversion model based on feature synthesis) (image source: generated by AI painting tool)
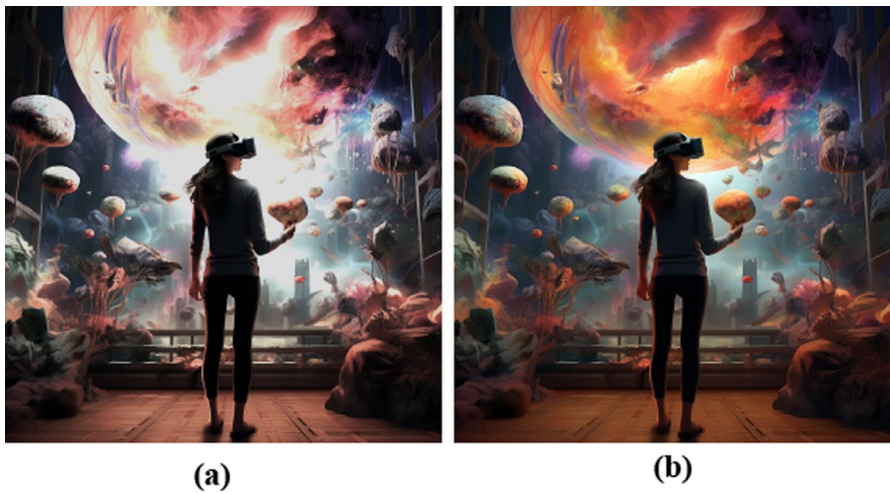

Fig. 9 The artistic display and comparison of the same new media art image generated by various models (**a** automatic image generation model based on the IoT; **b** image style conversion model based on feature synthesis) (image source: generated by AI painting tool)

originated in mid-twentieth century America, accentuates the artist's personal emotions and creativity over realism. Abstract expressionist paintings usually use large areas of color blocks, lines, shapes, and vibrant color contrasts to convey the artist's inner world. Impressionism is an artistic movement that rose in France at the end of the nineteenth century, emphasizing the capture of light and color, as well as the description of natural landscapes and daily life. Impressionist paintings frequently employ small, swift brushstrokes, bright, saturated colors, and often overlook details and outlines to express the artist's immediate impression of the scene. According to the above comparison, Fig. 8b, generated by the image style conversion model based on feature synthesis, is better suited to the thematic and atmospheric context of new media art. The impressionist style excels at reflecting the color and light variations among different images displayed on multiple screens in a gallery, while conveying the futuristic and dreamlike ambiance of the gallery itself. In contrast, the style of abstract expressionism proves overly weighty and somber, misaligned with the artworks showcased within the gallery's confines.

Figure 9a and b shows artistic transformations of the same image generated by various models. Figure 9a exhibits lower artistic quality, while Fig. 9b attains a higher level of artistic merit. In terms of color, figure b employs more vivid and saturated color, which makes the picture more vivid and attractive. Conversely, Fig. 9a favors darker, subdued colors, resulting in a more somber and lackluster visual impression. Regarding details, Fig. 9b features clearer, finely delineated elements, enhancing the image's overall vividness and richness. In contrast, Fig. 9a showcases vaguer, coarser details, contributing to a more indistinct and monotonous composition. From the contrast point of view, Fig. 9b presents a higher contrast of light, shade, and color, which makes the picture more prominent and hierarchical. Figure 9a offers diminished contrasts in light, shadow, and color, resulting in a flatter and less engaging image. From the aspect of composition, Fig. 9b adopts a better perspective and layout, which makes the picture more orderly and interesting. In contrast, Fig. 9a utilizes a less favorable perspective and layout, contributing to a more chaotic and uninspiring portrayal. From the theme, Fig. 9b expresses a stronger fantasy and imaginary theme, infusing the image with deeper meaning and interest. Figure 9a presents a weaker sense of fantasy and imagination, rendering the image less significant and engaging. Therefore, Fig. 9b, generated by the image style conversion model based on feature synthesis, exhibits greater artistic quality by demonstrating enhanced aesthetics and creativity in terms of color, detail, contrast, composition, and theme.

5.3 Discussion

The experimental data and evaluation method presented in this study are actually effective and reliable in evaluating the style transfer and image artistry of new media art images generated by different models. When comparing the image style transfer accuracy of M1–M6 models with various parameters, the recognition accuracy of M1-M6 models is higher than 83.51%. Further, compared with other model algorithms (such as CNN, AlexNet, etc.), the proposed model algorithm is significantly

higher, and the image recognition accuracy exceeds 96.03%, which is obviously superior to other algorithms. Meanwhile, the application prospect of AI and IoT technology in new media art creation is also discussed to further explore the feasibility and practicability of the model. In creating new media art, AI and IoT technology have broad application prospects. AI technology empowers artists to harness big data and algorithms to achieve unique effects, such as image generation, style transformation, and content creation. This affords artists greater creative possibilities. In parallel, the progression of IoT technology enables real-time interactions between new media artworks and their audiences. Utilizing sensors and network technologies, these works engage in two-way communication and audience participation. This integration has brought a more open and diverse creative platform for artistic creation, which can break the time and space restrictions in the virtual space and create rich and diverse artistic experiences. Through the application of AI and IoT technology, artworks can make real-time changes and adjustments according to the audience's needs and emotional feedback, making the interaction between art and audience closer and more personalized. This promising prospect expands the horizons and potential for developing new media art, stimulating discussions about the feasibility and practicality of the model.

In this study, it is mainly compared with the image works of Zhou et al. [40], Xu et al. [38], and Alexandru et al. [2]. In general, the images generated through parametric methods exhibit a greater degree of artistry compared to those created via nonparametric methods. The reason behind this distinction lies in the process of parameterized style transfer, where the desired target image effect is synthesized by constraining local content loss and global style loss. This objective aims to minimize the overall loss of the synthesized target image. However, it comes at the cost of diminishing the alignment of style features to some extent, resulting in an inadequate match of certain structural styles and textures. In contrast, the synthesis method used here is directly copying from the sampling of the feature map to the local feature information. By expressing the depth features of the image and accepting the better features of the domain, the obtained target image effect can keep more style and texture features. Compared to prior literature, the main contribution of this study is to optimize the creative mechanism of new media artistic features by using the method of layer-by-layer feature synthesis. In the existing image feature conversion method, the texture sampling method of feature points is single, and the result of the generated target image is monotonous. By comparing the other three schemes, the optimization method of feature points and the creation mechanism of new media art are improved, offering practical application value for enhancing image content and style conversion.

Comparing the results of this study with those of previous studies, Guo et al. [14] studied the application of cooperative neural networks in the IoT to detect powerful spammers. Using big data and various sensor data in the IoT, researchers established a cooperative neural network model and effectively detected the spammers by analyzing their behavior patterns [14]. Attwood [3] changed the social learning theory by relying on the IoT and AI. Researchers argued that the IoT and AI technology can offer new opportunities and ways for social learning, change how people acquire knowledge and communicate, and have a far-reaching impact on social

learning theory [3]. Misra et al. [24] studied the application of IoT, big data, and AI in the agriculture and food industries. Researchers pointed out that IoT technology can monitor and collect data in the process of agriculture and food production in real-time. Besides, big data and AI technology can analyze and process these data, provide intelligent decision support, and improve the efficiency and quality of the agriculture and food industry [24]. Grover et al. [13] reviewed the academic literature and social media discussions to deeply understand the application of AI in operation management. The researchers summarized some key insights, including the application fields, application effects, and challenges of AI technology in operation management, and prospected the future research direction [13].

To sum up, the theoretical and practical contribution of this study is that the constructed model successfully improves the accuracy of image recognition, which is of great significance to the field of new media art and image processing. High-accuracy image recognition helps improve the performance of automated image processing and recognition systems, providing art creators and practitioners with more powerful tools. By means of parameterization, the proposed model can generate more artistic images. This could promote innovation and development in the field of new media art, create a broader space and possibilities for the development of new media art, and enhance the interaction between artworks and audiences.

6 Conclusion

In the field of modern art, with the continuous progress of science and technology, sensors have become indispensable tools in artistic creation. Through the network's connection, artists can communicate with the audience in two directions and realize the participatory artistic experience. This open and diversified creative platform enables artworks to transcend the limitations of time and space and be displayed in virtual space. Integrating sensors and network connectivity allows the audience to interact with artworks and create a richer artistic experience. Artists can fully unleash their imagination and creativity in this new creative environment and create amazing artworks. Simultaneously, the audience has become more active and involved in the process of artistic creation. This new art form breaks the traditional time and space restrictions and brings new possibilities for artistic creation. With the development of Metaverse vision as the background, a new feature synthesis method of artistic images from deep to shallow is constructed by applying IoT technology based on visual geometry groups and AI. It combines the strengths of traditional texture synthesis and artistic style transmission techniques. In this study, the style transmission method based on feature synthesis is analyzed by visual geometry and texture synthesis technology. The model structure diagram based on shallow features is generated by obtaining feature images of different network levels, and the VGG19 network further synthesizes the virtual experimental images. The final comparative experimental results reveal that compared with other research results, the proposed style transfer method based on feature synthesis can restrain the unreasonable distribution of texture elements in traditional images. Consequently, it enhances both content and style expression within the generated images, ultimately yielding a more

profound artistic impact. Nevertheless, this study exhibits certain limitations. The primary constraint is the instability of image data classifiers in classification results. Subsequent research should aim to develop a real-time model for new media art creation and style transformation. This model should hinge on the style transformation of synthetic images, offering a practical reference for enhancing the mechanisms of new media art creation. Additionally, it should incorporate extensive classification data to bolster its effectiveness.

Data availability The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Ahmadkhani S, Moghaddam ME (2022) Image recommender system based on compact convolutional transformer image style recognition. J Electron Imaging 31(4):043054
2. Alexandru I, Nicula C, Prodan C et al (2022) Image style transfer via multi-style geometry warping. Appl Sci 12(12):6055
3. Attwood AI (2020) Changing social learning theory through reliance on the internet of things and artificial intelligence. J Sustain Soc Change 12(1):8
4. Berente N, Gu B, Recker J et al (2021) Managing artificial intelligence. MIS Q 45(3):1433–1450
5. Bourached A, Cann GH, Griffths RR et al (2021) Recovery of underdrawings and ghost-paintings via style transfer by deep convolutional neural networks: a digital tool for art scholars. Electron Imaging 2021(14):42–51
6. Castellano G, Vessio G (2021) Deep learning approaches to pattern extraction and recognition in paintings and drawings: an overview. Neural Comput Appl 33(19):12263–12282
7. Cvitić I, Perakovic D, Gupta BB et al (2021) Boosting-based DDoS detection in internet of things systems. IEEE Internet Things J 9(3):2109–2123
8. Dana LP, Salamzadeh A, Mortazavi S, Hadizadeh M (2022) Investigating the impact of international markets and new digital technologies on business innovation in emerging markets. Sustainability 14(2):983
9. Elemento O, Leslie C, Lundin J et al (2021) Artificial intelligence in cancer research, diagnosis and therapy. Nat Rev Cancer 21(12):747–752
10. Esenogho E, Djouani K, Kurien AM (2022) Integrating artificial intelligence Internet of Things and 5G for next-generation smartgrid: a survey of trends challenges and prospect. IEEE Access 10:4794–4831

11. Feuerriegel S, Shrestha YR, von Krogh G, Zhang C (2022) Bringing artificial intelligence to business management. Nature Machine Intell 4(7):611–613

12. Finlayson SG, Subbaswamy A, Singh K et al (2021) The clinician and dataset shift in artificial intelligence. N Engl J Med 385(3):283

13. Grover P, Kar AK, Dwivedi YK (2022) Understanding artificial intelligence adoption in operations management: insights from the review of academic literature and social media discussions. Ann Oper Res 308(1–2):177–213

14. Guo Z, Shen Y, Bashir AK, Imran M, Kumar N, Zhang D, Yu K (2020) Robust spammer detection using collaborative neural network in Internet-of-Things applications. IEEE Internet Things J 8(12):9549–9558

15. Gupta R, Srivastava D, Sahu M et al (2021) Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Mol Diversity 25(3):1315–1360

16. Hien NLH, Van Huy L, Van Hieu N (2021) Artwork style transfer model using deep learning approach. Cybern Phys 10(3):127–137

17. Huang L, Tei S, Wu Y et al (2021) A support system for artful design of tessellations drawing using CNN and CG. Int J Affect Eng 20(2):95–104

18. Jung S, Lee H, Myung S et al (2022) A crossbar array of magnetoresistive memory devices for in-memory computing. Nature 601(7892):211–216

19. Kang J, Lee S, Lee S (2021) Competitive learning of facial fitting and synthesis using uv energy. IEEE Trans Syst Man Cybern Syst 52(5):2858–2873

20. Kuzlu M, Fair C, Guler O (2021) Role of artificial intelligence in the Internet of Things (IoT) cybersecurity. Discov Internet Things 1(1):1–14

21. Liao YS, Huang CR (2022) Semantic context-aware image style transfer. IEEE Trans Image Process 31:1911–1923

22. Lombardi M, Pascale F, Santaniello D (2021) Internet of things: a general overview between architectures, protocols and applications. Information 12(2):87

23. Lyu Y, Lin CL, Lin PH et al (2021) The cognition of audience to artistic style transfer. Appl Sci 11(7):3290

24. Misra NN, Dixit Y, Al-Mallahi A, Bhullar MS, Upadhyay R, Martynenko A (2020) IoT, big data, and artificial intelligence in agriculture and food industry. IEEE Internet Things J 9(9):6305–6324

25. Mukundan A, Huang CC, Men TC et al (2022) Air pollution detection using a novel snap-shot hyperspectral imaging technique. Sensors 22(16):6231

26. Nasri A, Huang X (2022) Images enhancement of ancient mural painting of Bey's palace constantine, Algeria and Lacuna extraction using mahalanobis distance classification approach. Sensors 22(17):6643

27. Qazi EUH, Zia T, Almorjan A (2022) Deep learning-based digital image forgery detection system. Appl Sci 12(6):2851

28. Radanliev P, De Roure D (2023) New and emerging forms of data and technologies: literature and bibliometric review. Multimed Tools Appl 82(2):2887–2911

29. Radanliev P, De Roure D, Nicolescu R, Huth M, Santos O (2021) Artificial intelligence and the internet of things in industry 4.0. CCF Trans Pervasive Comput Interact 3:329–338

30. Ren W, Zhang J, Pan J et al (2021) Deblurring dynamic scenes via spatially varying recurrent neural networks. IEEE Trans Pattern Anal Mach Intell 44(8):3974–3987

31. Salem NM (2021) A survey on various image inpainting techniques. Future Eng J 2(2):1

32. Sharma G, Umapathy K, Krishnan S (2022) Trends in audio texture analysis, synthesis, and applications. J Audio Eng Soc 70(3):108–127

33. Shi Y, Yang K, Jiang T, Zhang J, Letaief KB (2020) Communication-efficient edge AI: algorithms and systems. IEEE Commun Surv Tutor 22(4):2167–2191

34. Suma V (2021) Internet-of-Things (IoT) based smart agriculture in India-an overview. J ISMAC 3(01):1–15

35. VALDEZ PF (2022) Art in the age of artificial intelligence: Japanese artistic painting style transfer through neural networks. Asian Stud 58(1):129–139

36. Verma B, Zarei O, Zhang S et al (2022) Development of texture mapping approaches for additively manufacturable surfaces. Chin J Mech Eng 35(1):1–14

37. Vulimiri PS, Deng H, Dugast F et al (2021) Integrating geometric data into topology optimization via neural style transfer. Materials 14(16):4551

38. Xu Z, Hou L, Zhang J (2022) IFFMStyle: high-quality image style transfer using invalid feature filter modules. Sensors 22(16):6134

39. Yamashita R, Long J, Banda S et al (2021) Learning domain-agnostic visual representation for computational pathology using medically-irrelevant style transfer augmentation. IEEE Trans Med Imaging 40(12):3945–3954
40. Zhou Y, Yu K, Wang M et al (2021) Speckle noise reduction for OCT images based on image style transfer and conditional GAN. IEEE J Biomed Health Inform 26(1):139–150

Authors and Affiliations

Xiaolong Wang[1] · Ling Cai[1] · Yunhao Xu[2]

✉ Xiaolong Wang
   soa_wangxl@ujn.edu.cn

   Ling Cai
   sa_cail@ujn.edu.cn

   Yunhao Xu
   P117579@siswa.ukm.edu.my

[1] School of Fine Arts, University of Jinan, Jinan 250022, Shandong, China

[2] National University of Malaysia, 43600 Bangi, Selangor, Malaysia