IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 4 July 2023, accepted 24 July 2023, date of publication 28 July 2023, date of current version 9 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3299863

RESEARCH ARTICLE

Three-Dimensional Animation Generation and Image Enhancement Technology Based on Multi-Column Convolutional Neural Network Model

WEIRAN CAO[1], YOUPING TEN[1], CHUNXUAN YU[2], JINGMIN SUN[1], AND JUN MAO[3]
[1]School of Art and Archaeology, Hangzhou City University, Hangzhou 310015, China
[2]GE Appliances, Louisville, KY 40225, USA
[3]College of Art Design, Zhejiang Gongshang University Hangzhou College of Commerce, Hangzhou 311599, China

Corresponding author: Jun Mao (192030011@zjhzcc.edu.cn)

This work was supported in part by the Major Humanities and Social Sciences Research Projects in Zhejiang Higher Education Institutions under Award 2023GH028.

ABSTRACT This paper aims to improve the quality and fidelity of three-dimensional (3D) animation. Firstly, the application model of Multi-Column Convolutional Neural Network (MCNN) in 3D animation generation and image enhancement is proposed. Aiming at the generation of 3D animation, the MCNN algorithm suitable for this field is selected, and its working principle is explained in detail. Meanwhile, the theoretical basis of 3D animation generation is introduced, which provides a theoretical basis for subsequent experiments. Secondly, for image enhancement, the MCNN is also selected as the key technology, and its application model in image enhancement is explained. Finally, a simulation experiment is carried out to evaluate the effect of the proposed MCNN model in 3D animation generation and image enhancement. By collecting appropriate data sets and setting parameters in the corresponding experimental environment, the performance of the proposed model is evaluated. The results show that, compared with the traditional methods, the MCNN model shows better performance and effect in animation generation and image enhancement tasks. Specifically, this method can still maintain good performance under the conditions of shorter training time, faster reasoning time and lower memory occupation, and this method has advantages in computational efficiency. 3D animation generation and image enhancement technology with MCNN model can significantly improve the animation quality and image fidelity, and satisfactory experimental results have been obtained. The experimental results in this paper verify the application potential of MCNN in 3D animation generation and image enhancement, and provide new ideas and directions for further research and application.

INDEX TERMS Multi-column convolutional neural network, three-dimensional animation, image enhancement, computational efficiency, image fidelity.

I. INTRODUCTION

A. RESEARCH BACKGROUND AND MOTIVATIONS

With the continuous development of science and technology and the expansion of application fields, three-dimensional (3D) animation generation and image enhancement technology play an important role in entertainment, advertising, virtual reality and other fields. 3D animation generation technology can simulate real-world objects and scenes, making cartoons, games and special effects more realistic [1]. Image enhancement technology can improve the image quality, enhance details, reduce noise, and improve the visual effect and information transmission ability of the image. The development and application of these technologies are very important to meet people's demand for high-quality visual experience [2]. However, there are still some challenges

The associate editor coordinating the review of this manuscript and approving it for publication was Chong Leong Gan.

and problems to be solved in the field of 3D animation generation and image enhancement. Traditional 3D animation generation methods often require complex modeling and rendering processes, which consume a lot of time and human resources [3]. Image enhancement technology also faces the problems of insufficient illumination, blurred image and noise interference [4]. Therefore, in order to improve the efficiency and quality of 3D animation generation and image enhancement, it is necessary to introduce new algorithms, models and technologies.

In this context, Multi-Column Convolutional Neural Network (MCNN) has become a research direction of great concern. MCNN is a deep learning model, which extracts features from images and videos through multiple convolution layers and pooling layers. It has the advantages of processing large-scale data, automatically learning feature representation and good generalization ability [5]. Therefore, the application of MCNN in 3D animation generation and image enhancement is expected to achieve better results and performance [6]. The motivation of this paper is to explore the application of MCNN in 3D animation generation and image enhancement. It is hoped to make use of the advantages of MCNN to improve the efficiency and quality of 3D animation generation, and improve the effect of image enhancement and the ability of detail preservation.

B. RESEARCH OBJECTIVES

The objectives of this paper include the following aspects: Firstly, the application of MCNN in 3D animation generation will be studied. By deeply understanding the theoretical basis of 3D animation generation, a MCNN model suitable for this field will be selected, and its principle and characteristics will be explained in detail. On this basis, a MCNN application model for 3D animation generation is proposed, aiming at improving the efficiency and quality of the generation process. Secondly, the application of MCNN in image enhancement is studied. The MCNN model suitable for image enhancement field is selected, and its working principle and characteristics are explained in detail. By introducing MCNN, a specific application model for 3D animation image enhancement is proposed, aiming at improving image quality, enhancing details and reducing noise [7].

Finally, simulation experiments are carried out to verify the performance of the proposed MCNN application model in 3D animation generation and image enhancement. The appropriate data sets are collected and simulated and evaluated in the experimental environment. The appropriate parameters are set, and the proposed models are objectively evaluated and compared through performance evaluation indicators. Through the analysis and discussion of experimental results, the effects and advantages of MCNN in 3D animation generation and image enhancement are evaluated, and suggestions for improvement and optimization are provided. By achieving the above research objectives, it is expected to provide useful contributions to the research and application of 3D animation generation and image enhancement. This paper helps to promote the development of these fields, and provides guidance and reference for academic research and practical application in related fields.

II. LITERATURE REVIEW

By using the powerful feature extraction ability and learning ability of Convolutive Neural Network (CNN), researchers have proposed a variety of 3D animation generation methods based on CNN, such as Anantrasirichai and Bull based on the method of generating confrontation networks [8], Kwolek et al. and the method of variational self-encoder, etc. [9]. These methods have achieved remarkable results in reducing manual participation, improving generation efficiency and enhancing generation quality. At present, image enhancement techniques are mainly divided into traditional methods and methods based on deep learning [10]. Traditional methods mainly include histogram equalization, filter and edge enhancement [11]. With the rise of deep learning, the image enhancement method based on deep learning has made remarkable progress. Especially, the application of CNN has brought a new breakthrough for image enhancement.

The CNN can better capture the features and context information of the image and improve the enhancement effect [12]. Some mainstream image enhancement methods based on deep learning include super-resolution reconstruction, image denoising and image restoration [13]. These methods realize the learning and recovery of local and global features of images by training deep neural networks with large-scale image data sets [14]. MCNN is a model based on CNN, which has multiple convolution layers and pooling layers [15]. In the task of image denoising, MCNN can learn the noise pattern in the image and improve the image quality through denoising operation [16]. In addition, Toshpulatov et al. applied the MCNN to image super-resolution reconstruction [17], and Hazourli et al. transformed the image style of MCNN into other image processing tasks [18], all of which achieved remarkable results.

Based on the review of related literature, this paper will further study and explore the application of MCNN in 3D animation generation and image enhancement. In the next section, the research model and method of this paper will be introduced in detail, and the experimental simulation will be carried out to verify the performance and effectiveness of the proposed method.

III. RESEARCH MODEL

A. MCNN IN 3D ANIMATION GENERATION

1) THEORETICAL BASIS OF 3D ANIMATION GENERATION

3D animation generation is a process of simulating a three-dimensional scene by computer technology and generating animation, as shown in Figure 1 [19].

Figure 1 involves several key concepts and technologies, such as 3D modeling, animation rendering, kinematics and key frame interpolation. Firstly, 3D modeling is the basis of

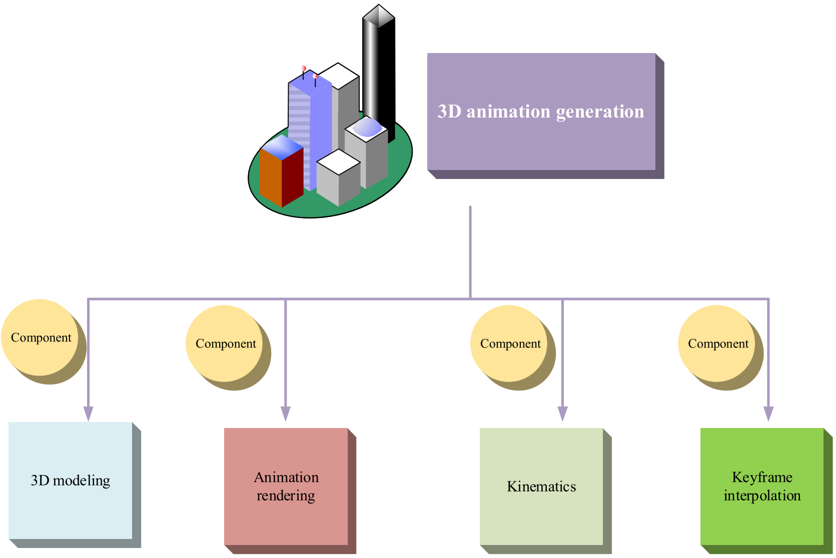

FIGURE 1. 3D animation generation.

3D animation generation, which involves modeling the geometric shape, texture and material properties of 3D objects. Secondly, animation rendering is the process of transforming the modeled 3D scene into the final image. In addition, kinematics is the description and simulation of object motion in animation. In 3D animation, the motion of objects can be realized by key frame interpolation [20].

2) APPLICATION MODEL OF MCNN IN 3D ANIMATION GENERATION

MCNN is a model based on convolutional neural network, which is widely used in 3D animation generation. The model structure of MCNN is shown in Figure 2 [21].

In Figure 2, MCNN extracts features from 3D scenes through multiple convolution layers and pooling layers, and trains them through back propagation algorithm to optimize and improve the animation generation process.

The application model of MCNN in 3D animation generation is as follows: Firstly, MCNN uses convolution layer to extract the features of 3D scene. The convolution layer performs convolution operation on the input feature map by sliding convolution kernel to obtain the output feature map. Each convolution kernel can learn different features. Multiple convolution layers can build a deep network structure, so that the network can learn more abstract and advanced feature representation. The process of convolution operation is shown in equation (1) [22].

$$Y(i,j) = \sum_{m} \sum_{n} X(i-m, j-n) \cdot K(m,n) \qquad (1)$$

$Y(i,j)$ is the pixel value of the output feature map. $X(i-m, j-n)$ is the pixel value of the input feature map. $K(m,n)$ is the weight of convolution kernel. m and n are index variables in convolution operation, which are used to indicate the position of convolution kernel when it slides on the input feature map.

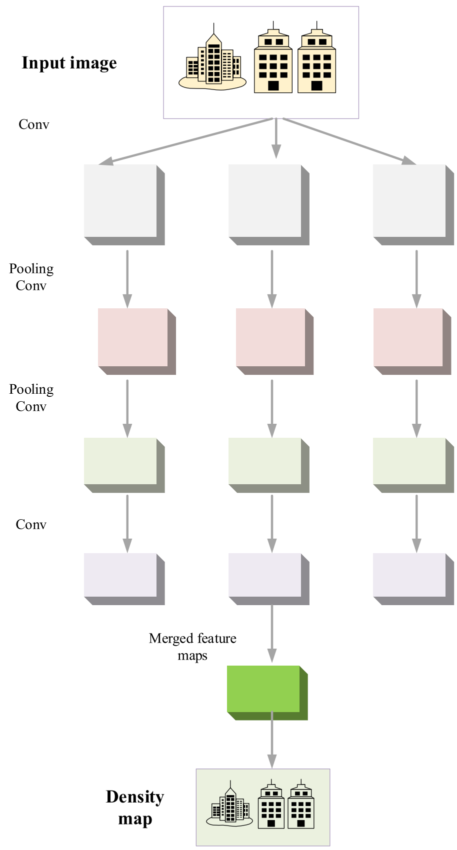

FIGURE 2. Model structure of MCNN.

Secondly, the pooling layer plays the role of down-sampling in the MCNN, which is used to reduce the size and parameters of the feature map, thus improving the calculation efficiency and reducing over-fitting [23]. The most commonly used pooling operations are average pooling and maximum pooling. Average pooling takes the average value of features in the pooled area as the output, while maximum pooling takes the maximum value as the output. The process of average pooling operation is shown in equation (2).

$$Y(i,j) = \frac{1}{k^2} \sum\nolimits_m \sum\nolimits_n X(k \cdot i + m, k \cdot j + n) \tag{2}$$

k is the size of the pool area.

In addition, activation function is often used to introduce nonlinear transformation in MCNN to increase the expressive power of the networks. The training process of MCNN usually adopts back propagation algorithm to optimize network parameters by minimizing loss function. Common loss functions include mean square error loss and cross entropy loss, and the loss function of mean square error loss is shown in equation (3).

$$L = \frac{1}{N} \sum\nolimits_{i=1}^{N} \left(Y_{\text{target}}^{(i)} - Y^{(i)}\right)^2 \tag{3}$$

L is the loss function value. N is the number of samples. $Y_{\text{target}}^{(i)}$ is the target output value. $Y^{(i)}$ is the predicted output value of the network.

In the training process, the back propagation algorithm is used to calculate the gradient and update the network parameters, so that the network can be gradually optimized and adjusted to achieve better animation generation effect.

In this paper, an innovative concept of adaptive convolution kernel is introduced to improve the quality and effect of 3D animation generation. Traditional MCNN usually use fixed convolution kernels to extract features in 3D animation generation. The innovation of this paper lies in the use of adaptive convolution kernel, whose weight can be dynamically adjusted according to the characteristics of input data [24]. The adaptive convolution kernel can be expressed by the following equation (4):

$$C(x,y) = \sum\nolimits_{i=1}^{m} \sum\nolimits_{j=1}^{n} W(i,j) \cdot I(x+i, y+j) \tag{4}$$

$C(x, y)$ represents the pixel value of the feature image of the convolution result. $I(x, y)$ represents the pixel value of the input image, and $W(i, j)$ represents the weight of the adaptive convolution kernel. Different from the traditional fixed convolution kernel, the weight $W(i, j)$ of the adaptive convolution kernel is automatically learned and adjusted according to the characteristics of the input data. In order to realize the adaptive convolution kernel, the attention module is introduced into the network, and the weight distribution adapted to different scenes and features is learned through training. By minimizing the difference between the generated image and the real image in the training process, the network can learn an adaptive convolution kernel suitable for the current input data.

In order to better capture the time sequence characteristics of animation sequences, a time sequence modeling mechanism is introduced into the MCNN. By combining time sequence information, the relationship between consecutive frames can be predicted more accurately, and a more coherent and natural animation sequence can be generated. The Recurrent Neural Network (RNN) structure is adopted to model the time series information. Long Short-Term Memory (LSTM) unit is used as the basic unit of RNN. The LSTM unit has mechanisms such as memory unit, forgetting gate, input gate and output gate, which can effectively deal with long-term dependencies. For the 3D animation generation task, the time sequence information is expressed as the feature vector of the input sequence, and then these features are processed through the LSTM layer [25]. The LSTM layer can capture the time series relationship in the input sequence and output a hidden state representing context information. This hidden state contains the information of past frames, so that the generation of current frame can be influenced by previous frames.

In order to further improve the quality of 3D animation generation, an innovative optimization strategy is adopted in the training process. Firstly, the attention mechanism is introduced to enhance the attention of the network to key features. Attention mechanism allows the network to dynamically adjust the weight of features in the generation process to better capture important details. Secondly, in order to enhance the generalization ability of the model, the domain adaptive method is adopted. 3D animation generation involves multiple scenes and animation styles, and these data often have different feature distributions. In order to adapt the model to different data domains, a domain adaptive loss function is introduced. This loss function enables the model to realize feature transfer between different scenes by minimizing the distribution difference between the source domain and the target domain [26]. Specifically, the maximum mean variance (MMD) is adopted as the objective function of domain adaptation. Based on the above optimization strategies, the objective function can be expressed as equation (5).

$$\begin{aligned} \min_{\theta} \sum\nolimits_{i=1}^{N} L_{\text{gen}}(x_i, y_i; \theta) + \lambda_1 \cdot \sum\nolimits_{i=1}^{N} L_{\text{att}}(x_i, y_i; \theta) \\ + \lambda_2 \cdot MMD(x_s, x_t) \end{aligned} \tag{5}$$

L_{att} is the loss function of attention mechanism. x_i and y_i are the input and target output respectively. θ is the parameter of the model, and x_s and x_t are the characteristics of the source domain and the target domain respectively.

Based on the above contents, the application model flow of MCNN in 3D animation generation in this paper is as follows: data preprocessing, converting 3D animation data into input format that can be used for training, and extracting key time sequence information and spatial features. Model architecture design uses deep convolution neural network architecture with adaptive convolution kernel. Time series information modeling uses cyclic neural network and LSTM network to model. Application of adaptive convolution kernel: Feature extraction and convolution operation. Deconvolution and up sampling. Result generation and optimization.

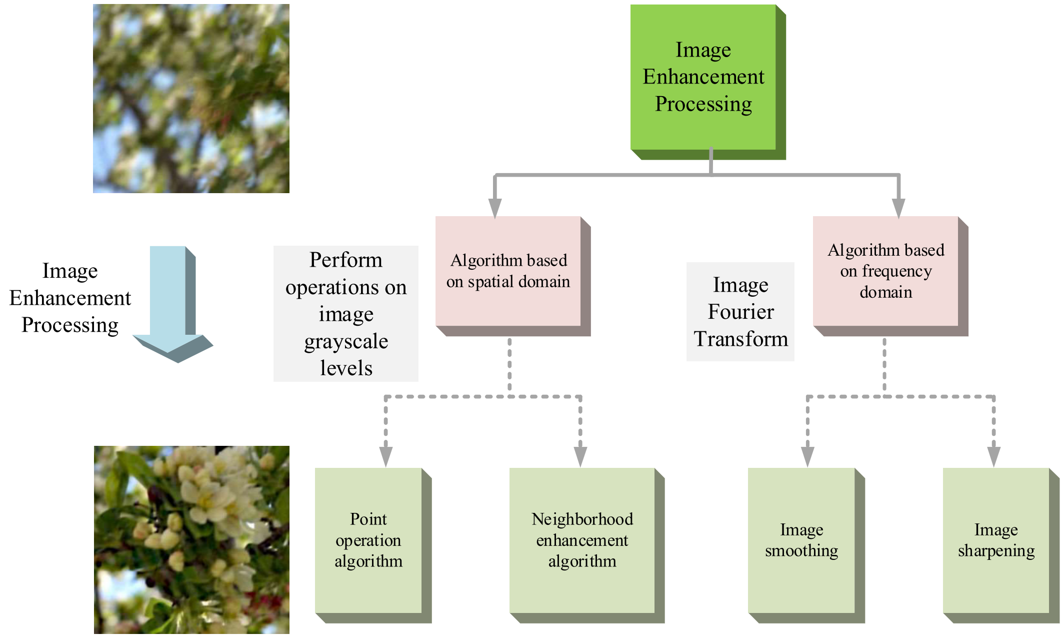

FIGURE 3. Content of image enhancement.

B. MCNN IN IMAGE ENHANCEMENT

1) THEORETICAL BASIS OF IMAGE ENHANCEMENT

Image enhancement is one of the key tasks to improve image quality and visual effect. The main contents of image enhancement are shown in Figure 3.

In Figure 3, the traditional image enhancement method is usually based on image processing technology and mathematical model. Among them, histogram equalization is one of the most commonly used traditional methods. In addition to histogram equalization, filter method is also widely used in image enhancement. Filter is a method to change the pixel value of an image by applying a specific filtering operation in the image. However, with the rise of deep learning, the image enhancement method based on deep learning has made remarkable progress. Deep learning model, especially CNN, shows great ability in image enhancement task. The image enhancement method based on deep learning enables the network to learn more advanced representation from the input image by designing appropriate loss function and network structure [27]. For example, the super-resolution reconstruction method uses the depth network to recover the details of high-resolution images from low-resolution images. Image denoising method learns noise model and image noise distribution by training network to realize image noise suppression.

2) APPLICATION MODEL OF MCNN IN 3D ANIMATION IMAGE ENHANCEMENT

Compared with the existing related research, this paper introduces the MCNN into the 3D animation image enhancement task, and adopts a specific model construction process to improve the enhancement effect and quality. The construction process of the model includes input and output definition, multi-layer convolution and pooling, activation function and batch normalization, residual connection and jump connection. The innovation lies in that for the 3D animation image enhancement task, the realism and detail retention ability of the enhanced image are improved through a well-designed model construction process [28].

First, define the input and output. The input is the original 3D animation image data, and the output is the enhanced image data. Next, multi-layer convolution and pooling operations are used to extract the features of the image. These operations are helpful to capture the local and global features of the image and provide valuable information for the subsequent enhancement steps.

In the design of the model, activation function and batch normalization operation are adopted to enhance the nonlinear ability and stability of the model, and the calculation process can be expressed as equation (6).

$$Z = \sigma(W \cdot X + b) \tag{6}$$

Z represents the feature map after activation function and batch normalization. σ represents the activation function. W represents the weight. X represents the input feature map. b stands for offset.

These operations help to keep the details of the image and improve the quality and visual effect of the enhanced image. In order to further improve the enhancement results, residual connection and jump connection are introduced. This connection can promote the flow and gradient spread of information and avoid the loss of information and the degradation of the model. By combining these connection methods, a more accurate and realistic enhanced image can be obtained.

In order to evaluate the performance of this paper, a set of suitable evaluation indexes is used. This includes the visual quality evaluation of the generated image, such as image clarity, texture details and fidelity. Commonly used evaluation indicators include Peak Signal to Noise Ratiom (PSNR) and Structural Similarity Index (SSIM) [29].

The calculation of PSNR is shown in equation (7).

$$PSNR = 10 * log10(MAX^2/MSE) \tag{7}$$

MAX is the maximum possible pixel value of the image, and MSE is the mean square error, which indicates the difference between the original image and the generated image.

The calculation of SSIM is shown in equation (8).

$$SSIM\ (x, y) = [l(x, y) * c(x, y) * s(x, y)]^{\alpha} \tag{8}$$

x and y represent the original image and the generated image respectively. $l(x, y)$ represents the similarity of brightness. $c(x, y)$ represents the similarity of contrast. $s\ (x, y)$ represents the similarity of structure, and α is the weight coefficient.

In the process of performance evaluation, a set of suitable experimental data sets and benchmark test images are used to ensure the objectivity and accuracy of the evaluation. Meanwhile, in order to verify the stability of the evaluation results, the experiment will be repeated many times and the experimental results will be statistically analyzed [30]. In order to verify the efficiency of the proposed method, the calculation time and resource consumption, including training time, reasoning time and memory occupation, are recorded. The experimental comparison is also made to compare the proposed MCNN method with other classic 3D animation generation and image enhancement methods. Compared with the existing methods, the advantages and improvements of the proposed method are evaluated.

IV. EXPERIMENTAL DESIGN AND PERFORMANCE EVALUATION

A. DATASETD COLLECTION

In order to carry out the experiment of 3D animation generation and image enhancement, this paper needs to collect appropriate data sets for training and testing. For 3D animation generation, the classic 3D model database ShapeNet is adopted, which contains abundant 3D model data. For image enhancement, Cifar-10, a widely used image data set, is selected, which contains all kinds of real-world images and covers different scenes and object categories.

B. EXPERIMENTAL ENVIRONMENT

The consistency and repeatability of the experimental environment is ensured when conducting experiments. A computer equipped with Nvidia GeForce RTX 2080 Ti graphics card is used as the experimental platform. The computer configuration is as follows: Intel Core i7 processor (8 cores, 3.6 GHz), 16GB (Gigabyte) memory and 1 TB (Terabyte) SSD (solid state drives) hard disk. The operating system is Windows 10, and an appropriate deep learning framework and library TensorFlow 2.0 are installed.

C. PARAMETERS SETTING

Parameter setting involves model selection, network structure, learning rate, batch size and so on. In the aspect of model selection, the method based on MCNN is adopted, and the specific network structure is adjusted according to the actual needs. In the design of network structure, factors such as the number of layers of the network, the size of convolution kernel and pooling mode are considered to achieve better generation effect.

Learning rate is an important parameter in the training process, which determines the speed of updating model parameters. Batch size is also a parameter that needs to be adjusted. Larger batch size can improve the efficiency of training, but it may lead to over-fitting of the model. Smaller batch size may lead to unstable training. There are some other parameters that need to be set, such as regularization parameters, selection of activation function, selection of optimizer, etc. Reasonable setting of these parameters can further improve the performance and generation quality of the model.

The parameter setting results in the two models in this paper are shown in Table 1.

TABLE 1. Parameter setting results.

Model	Generation of 3D images	Image enhancement
Network layer number	5 convolution layers and 3 fully connected layers (maximum pooling layer and convolution layer are used alternately)	3 convolution layers, 3 maximum pooling layers and 2 fully connected layers
Convolution kernel size	3x3	3x3
Pool mode	Maximum pooling	Maximum pooling
Learning rate	0.001	0.001
Batch size	16	64
Regularization parameter	0.001	0.0001
Activation function	ReLU	ReLU
Optimizer	Adam	Adam
Loss function	Mean square error loss	Mean square error loss

D. PERFORMANCE EVALUATION

In this paper, 300 data are randomly selected from the ShapeNet data set for the experiment, and the experiment was repeated three times. The performance evaluation result of 3D animation generation is shown in Figure 4.

Through the analysis of three groups of experimental data in Figure 4, it can be seen that the method in this paper is superior to the methods based onCNN and rules in terms of image clarity, texture details and fidelity. Specifically, the method in this paper shows the highest PSNR value in terms of image clarity, reaching 34.8, and the texture details are slightly reduced but still better than other methods, and the

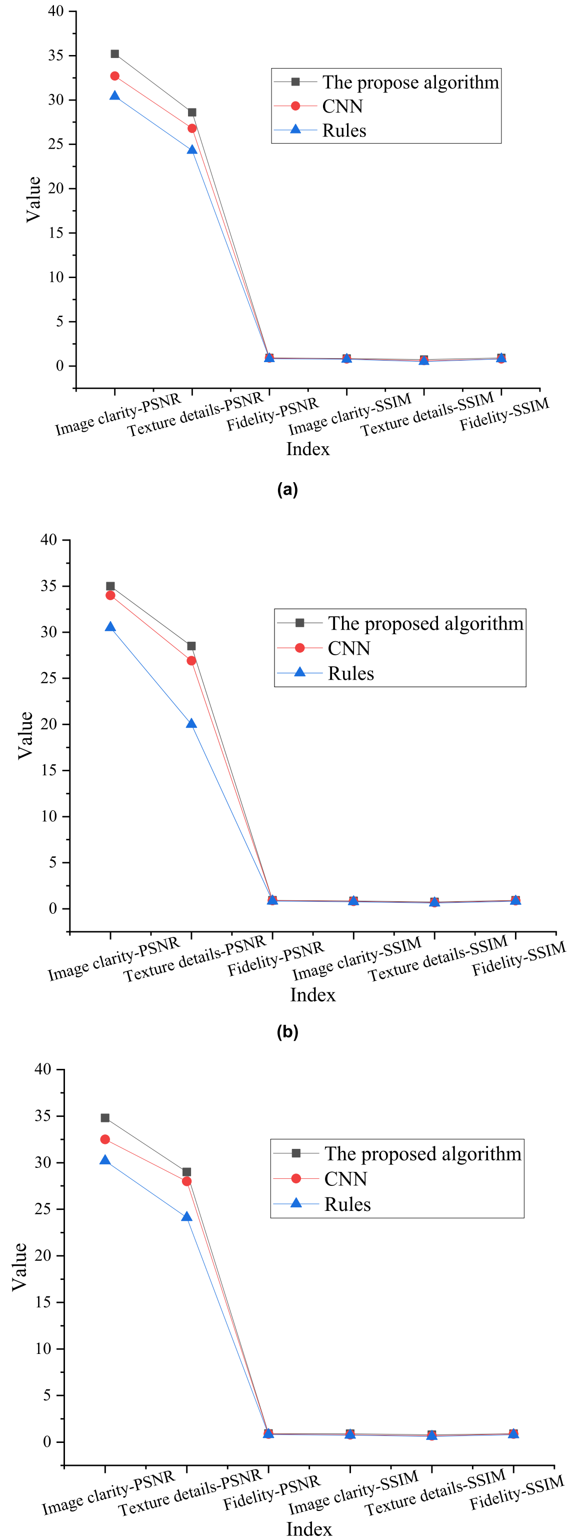

FIGURE 4. Performance evaluation results of 3D animation generation. (a) is the first experiment; (b) is the second experiment; (c) is the third experiment.

fidelity also shows the highest PSNR value. This shows that this method has obvious advantages in preserving details and improving image fidelity.

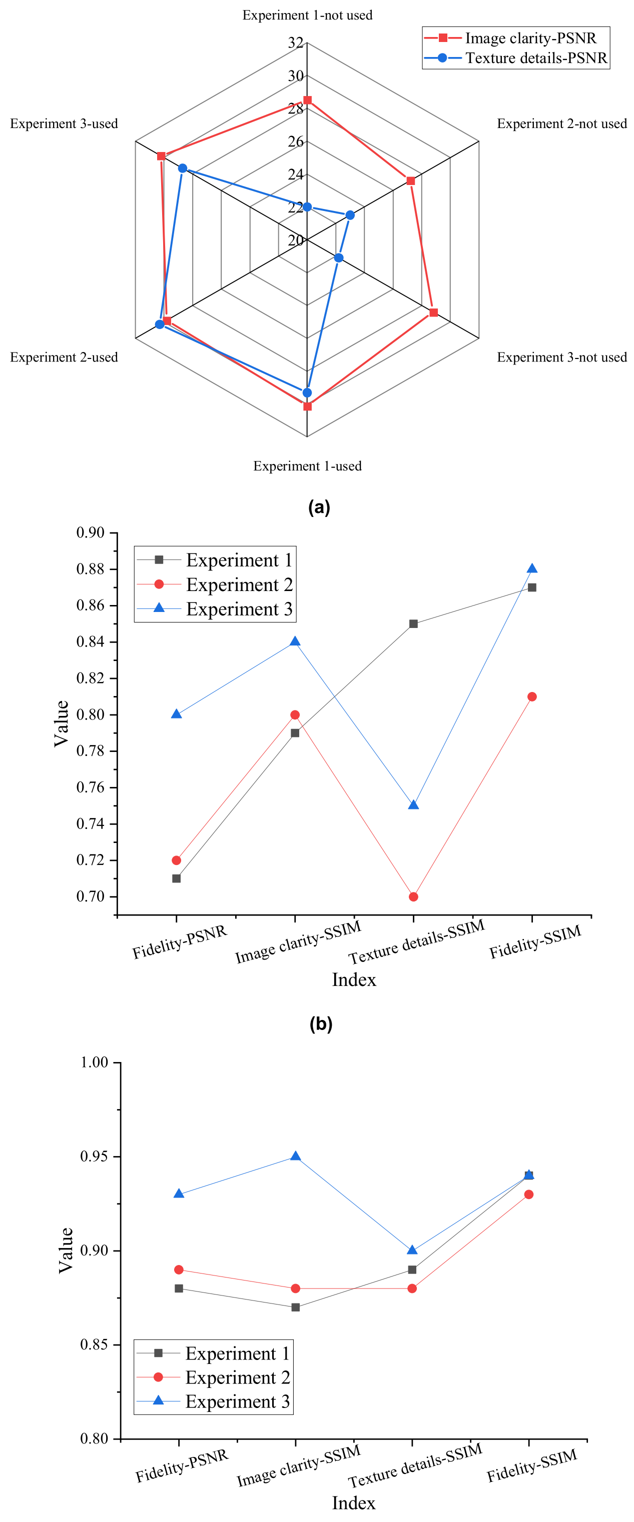

FIGURE 5. Performance evaluation results of image enhancement. (a) is the PSNR for clarity and PSNR for texture details; (b) is the SSIM of PSNR, clarity, texture details and fidelity without image enhancement; (c) is the PSNR for image enhancement and SSIM for clarity, texture details and fidelity.

In the same way, the performance evaluation result of image enhancement technology is shown in Figure 5.

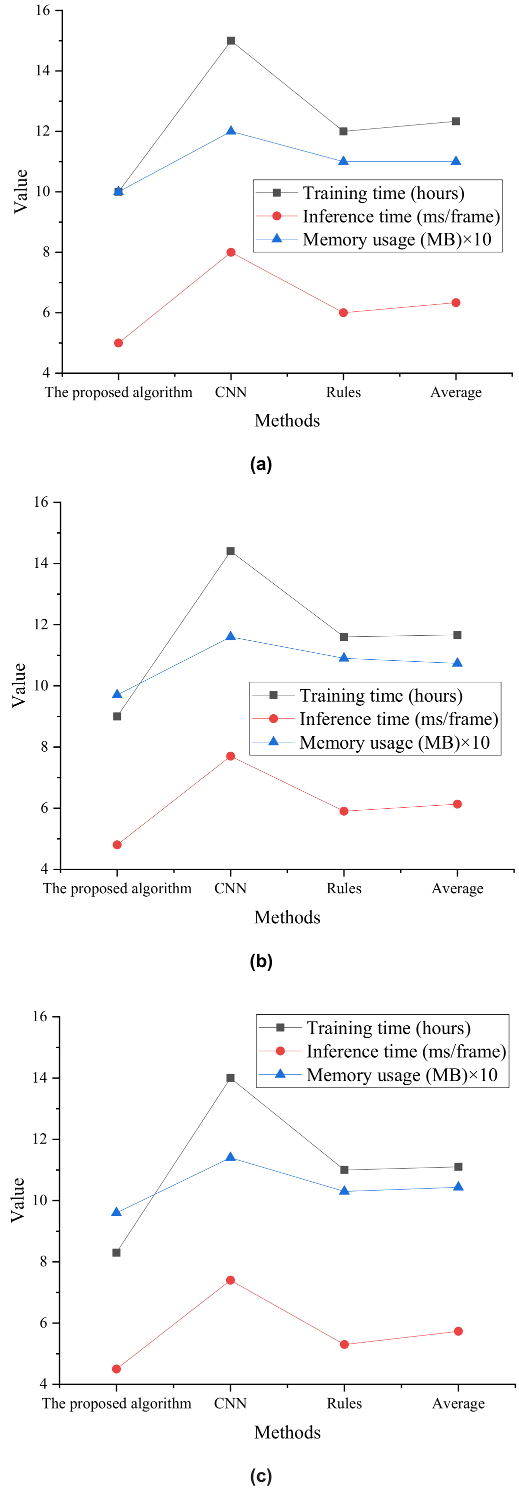

FIGURE 6. Calculation efficiency of 3D animation generation model. (a) is the first experiment; (b) is the second experiment; (c) is the third experiment.

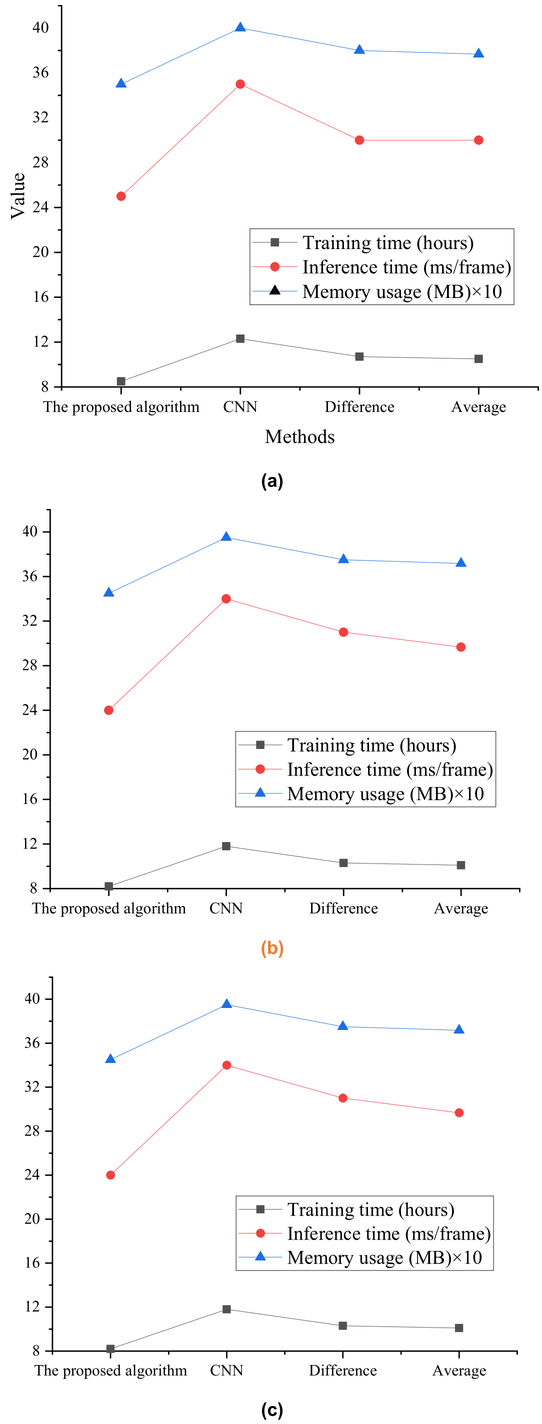

FIGURE 7. Calculation efficiency of image enhancement technology. (a) is the first experiment; (b) is the second experiment; (c) is the third experiment.

As can be seen from Figure 5, in terms of image clarity, the experimental results using image enhancement technology show that the average PSNR has increased from 28.1dB without image enhancement to about 30.0dB, which indicates that the reconstruction quality of the image has been significantly improved. Texture details and fidelity have also been improved, and the image enhancement technology proposed in this paper has achieved remarkable results in improving image quality.

The calculation efficiency results of the two methods are shown in Figure 6 and Figure 7.

As can be seen from Figures 6 and 7, the training time of the 3D animation generation model in this paper is about 9 hours, the reasoning time is at least 4.5 milliseconds per frame, and the memory occupation is at least 96MB (mbbyte). Compared with the methods based on convolutional neural network and rules, the 3D animation generation model in this paper is better. In the three experiments, the image enhancement technology in this paper also has better performance than the other two algorithms, the training time is at least 7.8 hours, the reasoning time is at least 23 milliseconds and the memory occupation is at least 340MB. Therefore, the method in this paper can still maintain good performance in the case of short training time, fast reasoning time and low memory occupation, and it has advantages in computational efficiency.

E. DISCUSSION

The method in this paper shows significant performance advantages in 3D animation generation and image enhancement, including better image clarity, texture detail preservation and fidelity improvement. These results show that the application of MCNN in 3D animation generation and image enhancement is effective, and it has broad application potential in these two fields. From the experimental results, this method shows the highest PSNR value in image clarity, which is due to the advantages of MCNN in capturing image features and preserving details. However, there is a slight decline in texture details, which may be due to some processing of the image in the process of image enhancement, resulting in some details blurred. Nevertheless, compared with other methods, this method still maintains a good ability to retain texture details. In addition, the experimental results also show that the use of image enhancement technology can significantly improve the image quality. By introducing the MCNN, the clarity of the image is successfully improved, and the texture details and fidelity of the image are enhanced. This shows that the application of MCNN in the field of image enhancement has potential, which can help improve the image quality and meet the needs of users for higher quality images. The research also pays attention to the computational efficiency of MCNN. Therefore, the method in this paper has not only made significant improvement in performance, but also has advantages in computational efficiency.

V. CONCLUSION

A. RESEARCH CONTRIBUTION

In this paper, the MCNN model is introduced in the field of 3D animation generation and image enhancement, and the following important research contributions are made. Firstly, the theoretical basis of 3D animation generation is explored, and the key problems and challenges in this field are analyzed. Secondly, the model of MCNN in 3D animation generation is designed and applied, and an effective algorithm or scheme is provided to generate realistic 3D animation. Meanwhile, the theoretical basis of image enhancement is also studied, and the MCNN is applied to 3D animation image enhancement to improve the quality and realism of the image. Experimental and simulation results verify the effectiveness and performance advantages of the proposed model. Therefore, this paper provides a new perspective and method for algorithms and technologies in the field of 3D animation generation and image enhancement, and has important guiding significance for further research and development in related fields.

B. FUTURE WORKS AND RESEARCH LIMITATIONS

However, although this method shows good performance in many aspects, there are still some limitations. First, the experimental data set is limited to specific scenes and image types, so the universality of the results may be limited. Secondly, this paper only focuses on the application of MCNN in 3D animation generation and image enhancement, while the application in other fields needs further exploration and research. Future research directions include expanding the scale and diversity of experimental data sets to verify the generalization ability and adaptability of this method. In addition, the structure and parameter settings of multi-convolution network can be further optimized to improve its performance and efficiency. In addition, people can consider comparing with other deep learning models or algorithms to explore their advantages and disadvantages in 3D animation generation and image enhancement.

REFERENCES

[1] Y. Xu, C. Jung, and Y. Chang, "Head pose estimation using deep neural networks and 3D point clouds," *Pattern Recognit.*, vol. 121, Jan. 2022, Art. no. 108210.

[2] L. Wang, "Cartoon-style image rendering transfer based on neural networks," *Comput. Intell. Neurosci.*, vol. 125, p. 1221, Feb. 2022.

[3] B. Mahmood, S. Han, and J. Seo, "Implementation experiments on convolutional neural network training using synthetic images for 3D pose estimation of an excavator on real images," *Autom. Construct.*, vol. 133, Jan. 2022, Art. no. 103996.

[4] S. Hossain, S. Umer, R. K. Rout, and M. Tanveer, "Fine-grained image analysis for facial expression recognition using deep convolutional neural networks with bilinear pooling," *Appl. Soft Comput.*, vol. 134, Feb. 2023, Art. no. 109997.

[5] K. Kumari, J. P. Singh, Y. K. Dwivedi, and N. P. Rana, "Multi-modal aggression identification using convolutional neural network and binary particle swarm optimization," *Future Gener. Comput. Syst.*, vol. 118, pp. 187–197, May 2021.

[6] H. Nguyen, T. Tran, Y. Wang, and Z. Wang, "Three-dimensional shape reconstruction from single-shot speckle image using deep convolutional neural networks," *Opt. Lasers Eng.*, vol. 143, Aug. 2021, Art. no. 106639.

[7] A. Sánchez-Caballero, S. de López-Diz, D. Fuentes-Jimenez, C. Losada-Gutiérrez, M. Marrón-Romera, D. Casillas-Pérez, and M. I. Sarker, "3DFCNN: Real-time action recognition using 3D deep neural networks with raw depth information," *Multimedia Tools Appl.*, vol. 81, no. 17, pp. 24119–24143, Jul. 2022.

[8] N. Anantrasirichai and D. Bull, "Artificial intelligence in the creative industries: A review," *Artif. Intell. Rev.*, vol. 19, pp. 1–68, May 2022.

[9] B. Kwolek, W. Baczynski, and S. Sako, "Recognition of JSL fingerspelling using deep convolutional neural networks," *Neurocomputing*, vol. 456, pp. 586–598, Oct. 2021.

[10] J. Lyu, H. Y. Lee, and H. Liu, "Color matching generation algorithm for animation characters based on convolutional neural network," *Comput. Intell. Neurosci.*, vol. 2022, p. 33, Dec. 2022.

[11] K. Saleh, A. Abobakr, M. Hossny, D. Nahavandi, J. Iskander, M. Attia, and S. Nahavandi, "Fast intent prediction of multi-cyclists in 3D point cloud data using deep neural networks," *Neurocomputing*, vol. 465, pp. 205–214, Nov. 2021.

[12] N. Jain, V. Gupta, S. Shubham, A. Madan, A. Chaudhary, and K. C. Santosh, "Understanding cartoon emotion using integrated deep neural network on large dataset," *Neural Comput. Appl.*, vol. 34, no. 24, pp. 21481–21501, Dec. 2022.

[13] A. K. Dubey and V. Jain, "An accurate recognition of facial expression by extended wavelet deep convolutional neural network," *Multimedia Tools Appl.*, vol. 81, no. 20, pp. 28295–28325, Aug. 2022.

[14] J. Ysique-Neciosup, N. Mercado-Chavez, and W. Ugarte, "DeepHistory: A convolutional neural network for automatic animation of museum paintings," *Comput. Animation Virtual Worlds*, vol. 33, no. 5, Sep. 2022, Art. no. e2110.

[15] H. Li, X. Jiang, B. Guan, R. Wang, and N. M. Thalmann, "Multistage spatio-temporal networks for robust sketch recognition," *IEEE Trans. Image Process.*, vol. 31, pp. 2683–2694, 2022.

[16] A. Aggarwal, M. Mittal, and G. Battineni, "Generative adversarial network: An overview of theory and applications," *Int. J. Inf. Manage. Data Insights*, vol. 1, no. 1, Apr. 2021, Art. no. 100004.

[17] M. Toshpulatov, W. Lee, and S. Lee, "Talking human face generation: A survey," *Expert Syst. Appl.*, vol. 219, Jun. 2023, Art. no. 119678.

[18] A. R. Hazourli, A. Djeghri, H. Salam, and A. Othmani, "Multi-facial patches aggregation network for facial expression recognition and facial regions contributions to emotion display," *Multimedia Tools Appl.*, vol. 80, no. 9, pp. 13639–13662, Apr. 2021.

[19] P. K. Shukla, M. Zakariah, W. A. Hatamleh, H. Tarazi, and B. Tiwari, "AI-DRIVEN novel approach for liver cancer screening and prediction using cascaded fully convolutional neural network," *J. Healthcare Eng.*, vol. 2022, p. 169, Dec. 2022.

[20] K. V. Dudekula, H. Syed, M. I. M. Basha, S. I. Swamykan, P. P. Kasaraneni, Y. V. P. Kumar, A. Flah, and A. T. Azar, "Convolutional neural network-based personalized program recommendation system for smart television users," *Sustainability*, vol. 15, no. 3, p. 2206, Jan. 2023.

[21] Y. Huo and S.-E. Yoon, "A survey on deep learning-based Monte Carlo denoising," *Comput. Vis. Media*, vol. 7, no. 2, pp. 169–185, Jun. 2021.

[22] S. A. Yazdan, R. Ahmad, N. Iqbal, A. Rizwan, A. N. Khan, and D.-H. Kim, "An efficient multi-scale convolutional neural network based multi-class brain MRI classification for SaMD," *Tomography*, vol. 8, no. 4, pp. 1905–1927, Jul. 2022.

[23] K. S. R. Prasad, P. Edukondalu, and G. S. Rao, "Comparative study and utilization of best deep learning algorithms for the image processing," *ResearchGate*, vol. 18, p. 227, Dec. 2022.

[24] A. Khan, S. Hayat, M. Ahmad, J. Cao, M. F. Tahir, A. Ullah, and M. S. Javed, "Learning-detailed 3D face reconstruction based on convolutional neural networks from a single image," *Neural Comput. Appl.*, vol. 33, no. 11, pp. 5951–5964, Jun. 2021.

[25] W. Bai, Z. Zhang, B. Li, P. Wang, Y. Li, C. Zhang, and W. Hu, "Robust texture-aware computer-generated image forensic: Benchmark and algorithm," *IEEE Trans. Image Process.*, vol. 30, pp. 8439–8453, 2021.

[26] Y. Yao, Z. Zhang, X. Ni, Z. Shen, L. Chen, and D. Xu, "CGNet: Detecting computer-generated images based on transfer learning with attention module," *Signal Process., Image Commun.*, vol. 105, Jul. 2022, Art. no. 116692.

[27] P. Mittal, "Retinal disease classification using convolutional neural networks algorithm," *Turcomat*, vol. 12, no. 11, pp. 5681–5689, Sep. 2021.

[28] M.-Y. Chen and H.-T. Wu, "Real-time intelligent image processing for the Internet of Things," *J. Real-Time Image Process.*, vol. 18, no. 4, pp. 997–998, Aug. 2021.

[29] I. Arent, F. P. Schmidt, M. Botsch, and V. Dürr, "Marker-less motion capture of insect locomotion with deep neural networks pre-trained on synthetic videos," *Frontiers Behav. Neurosci.*, vol. 15, Apr. 2021, Art. no. 637806.

[30] R. Liu, J. Shen, H. Wang, C. Chen, S.-C. Cheung, and V. K. Asari, "Enhanced 3D human pose estimation from videos by using attention-based neural network with dilated convolutions," *Int. J. Comput. Vis.*, vol. 129, no. 5, pp. 1596–1615, May 2021.