www.cad-journal.net Computer-Aided Design AND APPLICATIONS

New Media Art Interaction Design Based on Convolutional Neural Network

Huijuan Ren[1] and Xinchun Huang[2]

[1]Department of Communication Cultural Content, Cheongju University, Cheongju, Chungcheongbuk-do 361712, South Korea, rhuijuan2021@163.com
[2]Office of Science and Technology, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China, huangxinchun@nwpu.edu.cn

Corresponding author: Xinchun Huang, huangxinchun@nwpu.edu.cn

Abstract. New media art, as the product of the integration of science, technology and art, has unique charm. When people appreciate works of art, they will not only give priority to sensory experience, but also give priority to interactive experience. In the stage of new media art design, we should get the utmost out of modern sci & tech and computer aided design (CAD) software to combine new media with art design and improve the interactivity of new media art design. In order to get the utmost out of CAD technology to improve the interactivity of new media art, this article proposes a semantic segmentation algorithm for new media art images based on Atreus Convolutive Neural Network (ACNN, also called cavity convolutional neural network). The algorithm meets the requirements of different feature extraction by designing loss functions with different fusion degrees, and iteratively updates the reconstruction effect by using random gradient descent method to realize the fusion reconstruction of artistic image style and photo content. The test results show that the new media art image segmentation algorithm based on ACNN algorithm is better than the traditional convolutional neural network (CNN) algorithm in accuracy and efficiency. By retaining more spatial information and aggregating multi-scale feature information, this model can effectively segment the objects with different scales in the image and create a better artistic interaction experience in the art design of new media.

Keywords: New Media Art; Human-Computer Interaction; Atrous Convolution Neural Network; CAD
DOI: https://doi.org/10.14733/cadaps.2024.S7.270-283

1 INTRODUCTION

In traditional art, artists independently create works, and the audience can only passively accept the artist's spiritual ideas through the artistic images, and understand the artistic connotation in the works through imagination. Benjdira et al. [1] uses the generated adversarial Sexual network

(GAN) to unsupervised domain adaptation of aerial images, thus achieving semantic segmentation. Such as roads, buildings, trees, etc. Unsupervised domain adaptation refers to learning feature representations and classifiers from data without using label information, thereby achieving adaptation to different domains. In aerial images, different regions have different semantic information, such as airports, runways, aprons, buildings, lawns, etc. Using the generated adversarial Sexual network, each pixel in the aerial image can be classified into different categories, thus achieving the semantic segmentation of the aerial image. New media art design is based on interactivity, and realizes communication and interaction with the audience through information dissemination. When designing new media art, artists should pay special attention to interactive features, and join the artist's personal creation through modern sensing technology or imaging technology to jointly create works of art with strong interactive features. In the stage of new media art design, we should get the utmost out of modern sci & tech and CAD software to combine new media with art design and improve the interactivity of new media art design. New media works of art can effectively improve the transmission and utilization of information, and provide convenience and more beautiful things for people's lives. Therefore, it is needed to study the interactive characteristics in the art design of new media. In low light environments, due to insufficient light, the images captured by the camera are often very dim and contain a large amount of noise. To solve this problem, the algorithm first preprocesses the images of two channels and adopts multi-scale transformation fusion technology to fuse images at different scales to obtain richer information. Then, by performing weighted sparse representation on the image, important features in the image are effectively extracted and the impact of noise is reduced. Chen et al. [2] made sparse representation more accurate by weighting different parts of the image. Meanwhile, multi-scale transformation fusion technology can fuse image information at different scales, thereby increasing the amount of information and improving the clarity and contrast of the image. The combination of these two technologies has significantly improved the image processing performance of the algorithm in low light environments. Convolutional layer is one of the most basic computational layers in CNN, which extracts image features through convolutional operations. Convolutional operation is a mathematical operation that slides the convolutional kernel over an image and calculates the weighted sum of each position to generate a feature map. Ghosh et al. [3] Activation function maps the result of convolution operation to a higher dimensional space, thus increasing the expressiveness of the model. Common Activation function include ReLU, Sigmaid, Tanh, etc. The pooling layer is used to reduce the size of the feature map, reduce the number of model parameters, and thus improve the generalization ability of the model. Pooling operations typically use either maximum pooling or average pooling. CNN usually consists of multiple convolution layers, Activation function and pooling layers. By repeating the above steps, the deeper features of the image are gradually extracted. At the end of CNN, there is usually a fully connected layer that converts feature maps into prediction results, known as segmentation masks. CNN's training process uses the Loss function to measure the difference between the predicted results and the real results, and uses the optimizer to update the parameters of the model to minimize the Loss function. Through the above steps, CNN can learn the features in the image and use them to predict the segmentation mask of the image. This method has achieved good results in many applications, such as medical image segmentation, remote sensing image segmentation, etc.

Gu et al. [4] transformed the input feature map into a feature tensor with the same dimension through two convolutional layers and ReLU Activation function. Specifically, the spatial attention module will perform feature recalibration on each position in the input feature map, thereby enhancing the model's attention to different positions. Transfer the input characteristic graph to the first roll up layer and activate it with ReLU Activation function. Pass the result of the first step to the second roll up layer, and use ReLU Activation function to activate again. Pool the results of the second step to reduce the size of the feature map. Perform deconvolution on the pooled results to restore the size of the feature map. Multiply the deconvolution result with the original feature map one by one, and activate it with ReLU Activation function. Through these steps, the spatial attention module can learn the importance of different positions in the input feature map, thereby better focusing on important positions in subsequent tasks and improving the performance of the model. By synthesizing the weighted feature maps of these three modules, the final segmentation result can be obtained. At the same time, IA-CNN also provides an interpretable output that can display the confidence score and segmentation boundary of the segmentation results. Due to the growth of information generation and digitalization, traditional media gradually merged into new technologies such as digitalization and Internet, and thus new media was born. New media art is obviously different from traditional art design in the design process, but pays attention to interactive characteristics. Image is a method of expressing creative thinking in design, and it is the carrier of conveying design theme information and ideological content. In art design, different information content will have different expression images. Images use their vivid visual intuition to convey ideas and explain facts, which can make the audience understand the design content and accept the design information in a short time better than words. Images in design works can attract readers' attention and enhance the persuasiveness of design works. The art that needs to be expressed is expressed through CAD tools. The whole stage of art design is based on sci & tech. By using advanced technologies such as modern imaging technology and sensing technology, new works can be born through artists' creation. Because CAD technology has incomparable accuracy and aesthetic effect with traditional mode, its application in the field of art design has also become frequent. Art design software gradually replaces manual drawing tools, and art design through CAD technology has become an essential skill for contemporary designers. In this article, the optimization method of interactive design of new media art combined with CAD technology is studied, and the semantic segmentation of new media art images is carried out combined with ACNN to realize the fusion and reconstruction of art image style and photo content.

In the stage of new media art design, we should give priority to the interaction of new media art design and the formation of artistic character. Art design greatly enhances the space of art design and enriches the design skills. The core content of image semantic segmentation is the problem of image classification based on pixels, which mainly classifies every pixel of the image based on pixel expression, so as to achieve the semantic level division of the image. Because image semantic segmentation must meet the requirements of identification and location at the same time, that is, to determine the category of objects in the image and find the location of objects, it has become the most challenging task in computer vision technology. For the application of CAD technology in interactive design of new media art, this article has carried out the following research and innovation:

⊖ This article explores the connotation and value of interactivity in new media art design, understands the performance of interactivity in new media art design, and carries out semantic segmentation.

⊜ Starting from the core of CNN-convolution kernel, this article proposes an algorithm of fusing and reconstructing artistic images and photos with ACNN. The algorithm meets the requirements of different feature extraction by designing loss functions with different fusion degrees, and iteratively updates the reconstruction effect by using random gradient descent method to realize the fusion reconstruction of artistic image style and photo content.

This article analyzes a model for semantic segmentation of artistic images. The first section elaborates on the relevant background of new media art. In section 2, you have revised the citation of the results for relevant researchers and retrograde researchers. Section 3 provides an algorithmic explanation of interaction in media art design. Section 4 verifies the effectiveness and practicality of the proposed semantic segmentation method for new media art images. The results in Section 5 indicate that the new media art image segmentation algorithm based on ACNN algorithm has varying degrees of improvement in accuracy and efficiency compared to traditional CNN algorithms.

2 RELATED WORK

Gupta et al. [5] used genetic algorithms to optimize weights and then applied these weights to the image fusion process. By using genetic algorithms, this method can automatically adjust weights to produce high-quality HDR images. This method is widely used in many application fields, such as High dynamic range photography, night scene photography, low illumination photography, etc. By using genetic algorithms, this method can effectively process these challenging photography scenes and produce high-quality HDR images. In terms of pseudo exposure image fusion, this method can be applied to various HDR image synthesis tasks. The genetic algorithm method is a technology based on genetic algorithms that can effectively handle challenging photography scenes and produce high-quality HDR images. This method has been widely used in pseudo exposure image fusion and has shown good performance in many practical application scenarios. Hayat and Imran [6] use Recursive filter to filter the image of each exposure to obtain the best image details. By using different filters, this method can effectively balance images with different exposure levels, resulting in high-quality HDR images. This method is widely used in many application fields, such as High dynamic range photography, night scene photography, low illumination photography, etc. By using Recursive filter, this method can effectively process these challenging photographic scenes and produce high-quality HDR images. In multimedia tools, this method can be applied to various HDR image synthesis tasks, such as multi exposure image fusion, Tone mapping, etc. By using Recursive filter, this method can produce high quality HDR images and show good performance in many practical application scenarios. Based on deep learning technology, Khened et al. [7] preprocess the entire slide image, including image enhancement, data enhancement, image segmentation, and other operations. Convolutional neural network is a typical feedforward neural network, which can learn image features independently, and classify and predict. In Convolutional neural network, the main calculation is done by convolution layer, which extracts features by convolution operation on the input image. For feature extraction of full slide images, pre trained Convolutional neural network models, such as VGG, ResNet, Inception, can be used to fine tune these models to specific tasks. These models have achieved good performance in many image classification tasks, and can therefore be used to extract image features. After extracting the features of the full slide image, Convolutional neural network can be used to segment and predict the image. For segmentation tasks, you can use the convolution layer, pooling layer, deconvolution layer and other operations in the Convolutional neural network to convert the feature map into prediction results. For prediction tasks, a fully connected layer can be used to convert feature maps into prediction results. And each pixel in the image is classified into different categories. For example, cells, tissues, organs, etc. Analyze the segmented image, including morphological analysis, tissue analysis, cell counting, and other operations to obtain various information and features of the sample. Overall, the generalized deep learning framework for full slide image segmentation and analysis is a powerful image processing tool with broad application prospects in biomedical research, clinical diagnosis, disease analysis, and other fields. Adaptive morphological reconstruction image processing for seed image segmentation is an image processing method based on morphological reconstruction technology, used for seed image segmentation and processing. Morphological reconstruction technology is an image processing technology based on morphological operations, which can effectively remove noise and breakpoints in images, and enhance the continuity and connectivity of images. In seed image segmentation, Lei et al. [8] analyzed that morphological reconstruction techniques can remove noise and breakpoints from seed images while preserving their morphology and features, thus obtaining more accurate and complete seed images. Adaptive morphological reconstruction technology adaptively adjusts morphological operation parameters based on local features of an image, in order to better adapt to local features and changes in the image. In seed image segmentation, adaptive morphological reconstruction technology can adaptively adjust morphological operation parameters based on local features and changes in the seed image. In order to better remove noise and breakpoints, and enhance the continuity and connectivity of seeds.

Ma et al. [9] conducted image segmentation analysis using the technique of obtaining high-resolution radar images through synthetic aperture. The network can be optimized by Backpropagation to minimize the value of Loss function. After optimization, the network can automatically segment large SAR image data and output segmentation results. The use of attention graph convolutional layers for adaptive weighting and extraction of image features in different regions improves the accuracy and stability of segmentation. Large SAR image data typically has characteristics such as high resolution, high contrast, and high noise, which makes image segmentation of large SAR images challenging. Convolutional network is an image segmentation technique based on deep learning, which can adaptively learn and extract image features from different regions, thereby obtaining more accurate segmentation results. Merianos and Mitianoudis [10] use learning analysis transform to perform. By using learning analysis transformation, this method can generate high-quality HDR images and demonstrate good performance in many practical application scenarios. In summary, multi exposure image fusion using learning analysis transformation for HDR image synthesis is a deep learning based technique that can effectively handle challenging photography scenes and produce high-quality HDR images. This method has been widely used in imaging journals and has shown good performance in many practical application scenarios. Deep learning technology can automatically learn image features and process large-scale data. The subdivision method based on recurrent neural networks is a deep learning method that uses recurrent neural networks to segment and classify images. Different from the segmentation method based on Convolutional neural network, the subdivision method based on recurrent neural network can better deal with complex structure and shape images, so it has also been applied in medical image segmentation. Müller and Kramer [11] transform medical images into formats suitable for deep learning model processing. Convolutional neural network and other depth learning techniques are used to extract image features. Use deep learning models for segmentation prediction and generate segmentation results. Perform interpretability analysis on the segmentation results, display the confidence score and segmentation boundary of the segmentation results. In a word, Convolutional neural network and deep learning play an important role in medical image segmentation, which can improve the accuracy and interpretability of segmentation and provide more accurate support for medical decision-making. The fast HDR image generation method based on frequency division multiplexing technology for single snapshot images is a technique that achieves image fusion by dividing the spectrum into multiple sub spectra. Niu et al. [12] decomposed the spectrum of the input image into multiple sub spectra, and then processed each sub spectrum separately to obtain the best image details. By using the spectral differences between photos with different exposure levels, this method can effectively balance between photos with different exposure levels, resulting in high-quality HDR images. This method is widely used in many application fields, such as High dynamic range photography, night scene photography, low illumination photography, etc. In summary, the fast HDR image generation method based on frequency division multiplexing technology for single snapshot images is an effective technique. This can effectively balance photos with different exposure levels and produce high-quality HDR images. Qi et al. [13] preprocessed the input multi exposure images, including denoising, enhancement, and other operations, in order to better perform subsequent fusion operations. This method uses pixel level feature extraction algorithms, such as Local binary patterns (LBP) algorithm, to extract pixel level features of each exposure image. These features can be used to describe the local details of each image, such as texture, color, etc. Next, the method uses fusion rules to combine these features into the final HDR image. Semantic nighttime image segmentation is an important issue in the field of computer vision, which aims to segment different objects or regions in nighttime images and provide semantic interpretation.

In nighttime images, due to poor lighting conditions, the contrast and clarity of the image are poor, which poses a great challenge to semantic segmentation. Map guided course domain adaptation is a method that utilizes annotated data to train models, enabling them to adapt to different scenarios and tasks. In semantic nighttime image segmentation, map guided course domain adaptation methods can be used. Sakaridis et al. [14] used annotated map data to train the model and improve its segmentation accuracy in nighttime images. Uncertainty perception assessment is a method that can help models better understand the uncertainty and changes in images. In semantic nighttime image segmentation, there is significant uncertainty and variation in the image due to changes in lighting conditions and scenes. The use of uncertainty perception evaluation can help the model better understand these uncertainties and changes, thereby improving segmentation accuracy. Two-dimensional image semantic segmentation is widely used in many application fields, such as remote sensing image analysis, Medical imaging diagnosis, automatic driving, etc. In these applications, it is necessary to conduct in-depth analysis and understanding of images in order to extract useful information and provide support for decision-making. Typically, semantic segmentation of two-dimensional images is achieved through deep learning techniques. Common models include Convolutional neural network models (such as U-Net, SegNet, etc.), recurrent neural network models (such as CRF, etc.), and the generation of adversarial Sexual network models (such as Pix2Pix, etc.). These models can automatically learn the features of images and segment them into different regions, each with similar features. Ulku and Akagündüz [15] classified and summarized existing two-dimensional image semantic segmentation methods. Among them, CNN based methods are one of the most commonly used methods, with the basic idea of extracting image features through convolution operations and then using these features for segmentation. In addition, the method based on superpixel segmentation is also a commonly used method. The basic idea is to divide the image into a group of superpixels with similar colors, textures, and brightness, and then segment the superpixels. Yin et al. [16] fusing on image challenging problem. Traditional image fusion methods are often affected by factors such as changes in lighting and narrow dynamic range, making it difficult to obtain high-quality fusion results. To address this issue, researchers have proposed an image fusion method based on deep learning. These methods learn fusion strategies by training deep neural networks, which can automatically learn and optimize the fusion process, thereby achieving better fusion results. By using Reinforcement learning technology to train the deep neural network, this method can better deal with light changes and dynamic range. And it can generate high-quality intermediate images to achieve better fusion results. Depth stack transform is a depth learning technology based on Convolutional neural network, which can be used for medical image segmentation. It extracts image features by stacking multiple convolutional and pooling layers, and uses fully connected layers for classification. In medical image segmentation, deep stacking transform can be used for segmentation of various tissues and organs, such as lungs, liver, kidneys, etc. Different tissues and organs require different network structures for training and prediction. The advantage of deep stacking transformation is that it can automatically learn image features by stacking multiple convolutional and pooling layers, thereby obtaining better segmentation results. At the same time, the depth stacking transformation can also be trained using the Unsupervised learning method, thus avoiding the dependence on the labeled data. Zhang et al. [17] used cross validation, Confusion matrix and other methods to evaluate. To extend the deep learning technology of medical image segmentation to invisible fields, various factors need to be considered. And select and optimize according to specific application scenarios to obtain better segmentation results.

3 INTERACTIVITY IN NEW MEDIA ART DESIGN

Interaction is the most basic feature of new media art design, which has advantages that traditional art forms do not have. The interactivity of new media art design is mainly manifested through online media, which transmits information to the audience, communicates and interacts with the audience, and realizes the innovation of new media art forms. Nowadays, artistic design pays special attention to visual feeling. In new media art, sensory impact often plays a vital role in determining the quality of works. In the stage of creation, artists can effectively promote the interest of artistic works and increase their connotation by fully expressing visual impact. Interactive new media art is a new way of art communication in the new era. Some commonly used image devices let visitors see what artists want to express more intuitively and clearly, and through sensors and other devices, visitors naturally participate in the interactive process. Compared with the traditional media web design, the new media web design is rich in content, which effectively increases the efficiency of information updating, enables timely interaction and improves the effect of information dissemination. As far as concept is concerned, the new media art design is more interactive, which enables visitors to participate. The interactive features make visitors more selective and more humane than traditional media. As far as technology is concerned, the new media art design is based on the latest sci & tech, and combines with computers to create artistic works.

In the new media environment, the traditional linear communication mode is gradually replaced, and the audience actively participates in the stage of information dissemination, and the interaction of information dissemination is enhanced. Through IT, the new media art integrates the appreciator into the art, changes the traditional single sensory feedback, and enriches the appreciator's feelings from hearing, vision, touch and even taste. The growth of technology provides more design methods for art design, and this method is gradually recognized by people, forming a new aesthetic view. New media art design creates brand-new works of art with brand-new artistic forms, wonderful ideas and brand-new technologies. Visitors will transform their works in different ways, such as touching or sounding. The connection of new media art design can connect visitors all over the world. The interactivity of network art is mainly reflected in that people can control information freely and find information that can meet their own needs according to their own requirements in the network environment, thus saving information acquisition time and improving information dissemination effect.

New media art makes visitors change from traditional outsiders to participants in works of art. Because people are different, different people will have different experiences in the stage of participating in the creation of artistic works. To some extent, this can promote the innovation and growth of new media art and bring artists and visitors closer. The artistic interaction of new media has obvious aesthetic characteristics, and its aesthetic value is constantly improved through artistic design. At the present stage, in the stage of new media art design, the creator constantly carries out artistic conception through advanced sci & tech, and innovates artistic design on the basis of innovative thinking, so that the new media art design has the effect of "multiple media experiences" and realizes the transformation of different functions.

4 SEMANTIC SEGMENTATION MODEL OF NEW MEDIA ART IMAGES

When artists create new media art, they integrate the interactivity of IT into the works of art to the two-way interaction of hearing and touch, which greatly enriches the experience of the public participating in art design and creates conditions for the renewal of people's aesthetic concepts. This article will use ACNN to optimize the interactive design stage of new media art. Traditional methods can't achieve good results in image semantic segmentation, and graph theory is generally used as the pretreatment of image segmentation, while conditional random field is generally used as the follow-up work of image segmentation to improve the results. Then combined with the classification calculated by artificial feature extraction method, the classification of each pixel of the image by the two methods is obtained, and then the results are combined and optimized by conditional random field to get the final classification. The large receptive field enables convolution to extract features in a larger range. In image semantic segmentation, images need to rely on global information for feature extraction, and hole convolution is a good choice.

ACNN firstly convolves the neural network with the idea of CNN, and then replaces the convolution layer and pool layer in the network with the hole convolution layer and hole pool layer to ensure that the size of the feature map obtained by each layer is consistent with the original map. Finally, the pixel-level classification is completed through the convolution full connection layer, and the image semantic segmentation result consistent with the original image size is directly obtained. Its basic principle is shown in Figure 1.

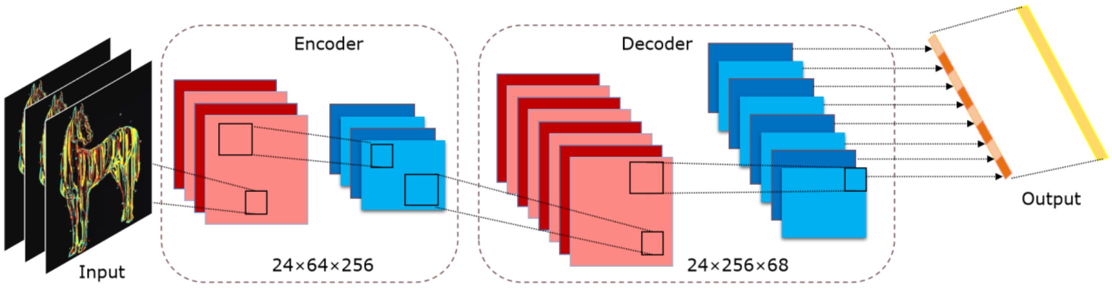

Figure 1: Basic principle of CNN.

In convolution operation, it can be understood as turning a 2D graph into a one-dimensional vector. The forward propagation of convolution operation can be understood as the matrix parameters of convolution kernel are left multiplied by this one-dimensional vector to get the output, and the backward propagation is the transposed left multiplied gradient vector of parameter matrix; The forward propagation of transposed convolution operation is the transposed left multiplication gradient vector of matrix parameters, which is the inverse operation of convolution operation, so it can be understood that the reverse propagation of convolution operation corresponds to the forward propagation of transposed convolution operation.

Convolution is the most important part of ACNN, which is used to extract feature information from data. The discrete convolution process is as follows:

$$H(x,y)=A\otimes k(x,y)=\sum_{M,N}A(m,n)k(x-m,y-n) \tag{1}$$

Assuming that the l-th layer is a fully connected layer, the weight matrix is W^l and the bias is b^l. The calculation stage of the fully connected layer is:

$$Z_j^l=f\left(W^l X^{l-1}+b^l\right) \tag{2}$$

Because the expressive ability of the linear model is not enough, in order to obtain the nonlinear model, a nonlinear activation function is added to the neural network to generate a nonlinear response to the input. Use ReLu as activation function:

$$f(x)=\max(0,x) \tag{3}$$

ReLu is a linear function, so its calculation speed is faster, which will make the network training speed faster.

The most important component of convolution layer is convolution kernel, and each convolution layer contains multiple trainable convolution kernels. Convolution kernel is actually a feature extractor, which can automatically extract features from the input image data. Different convolution kernel sizes will lead to different extracted feature information. In transposition convolution, the resolution of the image can be restored as long as the step size is set. The CNN lightweight method based on model compression does not redesign the model structure, but compresses the model from the parameter level on the basis of the existing model, so that the calculation and parameters of the model are reduced. ACNN convolution process is shown in figure 2.

The model constructed records the positional relationship between pixels through additional parameters in the down-sampling process, that is, the index of pooled layer. This enables the pixel information discarded due to the pooling operation in the up-sampling to be effectively restored by the recorded position information. In this article, the structural characteristics of hole convolution sampling.

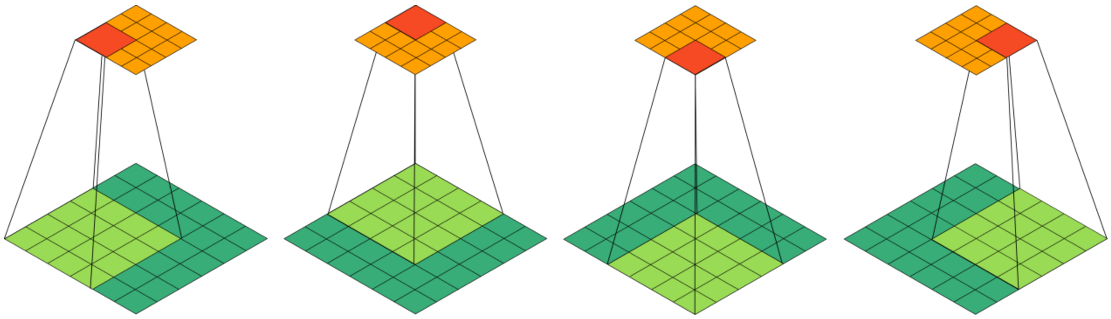

Figure 2: ACNN convolution process.

The hole convolution is simply applied to the residual network structure to achieve the purpose of lightweight, and further combined with point-by-point convolution to improve the lightweight effect, forming an improved hole convolution lightweight method.

Random pooling is to randomly select the pixel values in each sub-block, which can solve the problem of over-fitting in the training process:

$$p_i = \frac{a_i}{\sum_{k \in R_j}(a_k)} \tag{4}$$

Firstly, the probability p of each element of R_j is calculated, and then multiple distributed sampling operations are performed according to p. In order to avoid this situation, when using full connection, we usually do the Dropout operation:

$$h_{w,b}(x) = f\left(W^T x + b\right) \tag{5}$$

Where $h_{w,b}(x)$ represents output data, x represents input data, $W^T x$ represents connection weight, b represents offset, and $f(\ )$ represents activation function.

After convolution, the output of convolution layer can only be obtained through the operation of activation function, which is to impose nonlinear factors on the convolution results and improve the expression ability of the model.

$$\bigcup_{i=1}^{N} R_i = R \tag{6}$$

For all i and j, $i \neq j$, there are:

$$R_i \cap R_j = \varnothing \tag{7}$$

For $i = 1,2,\ldots,N$, there are:

$$P(R_i) = TRUE \tag{8}$$

For $i \neq j$, there are:

$$P(R_i \cup R_j) = FALSE \tag{9}$$

The advantage of convolution layer lies in local connection and weight sharing. The function expression of interactive design state model of new media art is given:

$$\lambda(HI) = Bt^{C*HI} \tag{10}$$

$$Z_i' = \sigma\left(\sum_{v_j \in N(v_i)} \tilde{\Delta}_{sym}[i, j](WZ_j)\right) \tag{11}$$

The model integrates useful semantic information and spatial edge information in different network layers through multi-scale feature fusion modules from shallow to deep and then from deep to shallow. In the decoding stage, the newly designed multi-task decoding network is used to help the multi-scale feature fusion module filter out useless noise information more accurately by pre-outputting supervised learning edge features and semantic features. In the final segmentation stage, the segmentation results are further refined by merging the learned semantic features and edge features to improve the overall performance of the network.

5 RESULT ANALYSIS AND DISCUSSION

In order to verify the effectiveness and practicability of the new media art image semantic segmentation method in this article, this section tests and analyzes the comprehensive performance of the algorithm. The experiment adopts MIT Scene Parsing Data set, which includes image semantic segmentation and image instance segmentation. A total of 150 categories are included, including the sky, roads, grasslands and discrete objects such as people, cars and beds. The training set in the data set contains 20,110 images, and the test set contains 2,200 images. Test the time required for image semantic segmentation with different numbers of pictures and different numbers of nodes, as shown in Figure 3.

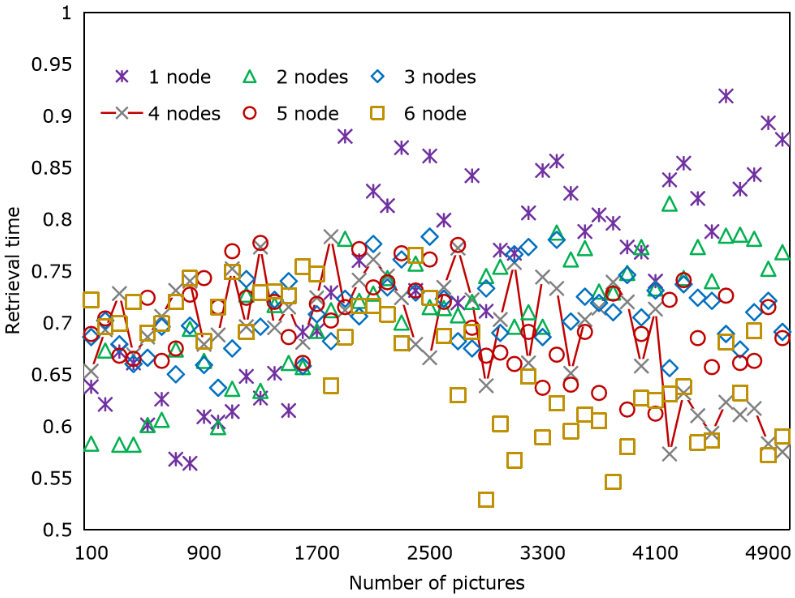

Figure 3: Image recognition consumes time.

In order to reduce the attenuation of accuracy, an improved hole convolution is combined with ordinary convolution, and a lightweight method of integrated hole convolution is proposed. As can be seen from Figure 3, when the quantity of artistic images is small, the more nodes there are, the image recognition efficiency may be negatively affected. With the increasing quantity of images, the efficiency of multiple nodes presents obvious advantages. The results of accuracy test using traditional CNN and ACNN algorithm are shown in Figure 4 and Figure 5.

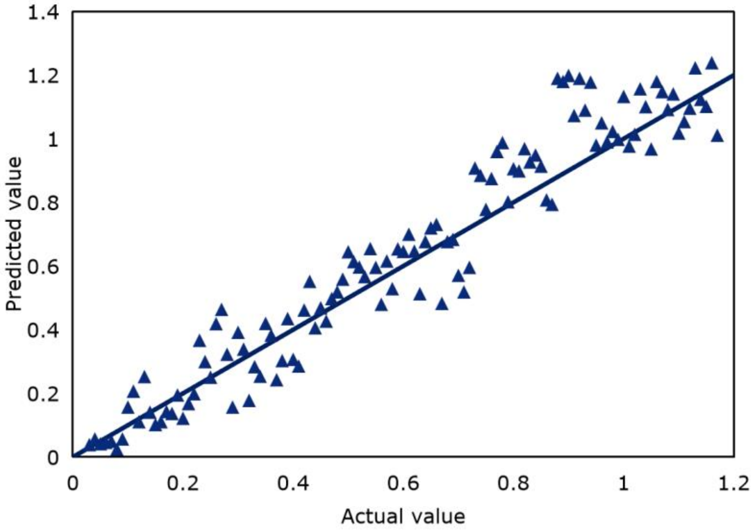

Figure 4: Accuracy test results of traditional CNN algorithm.

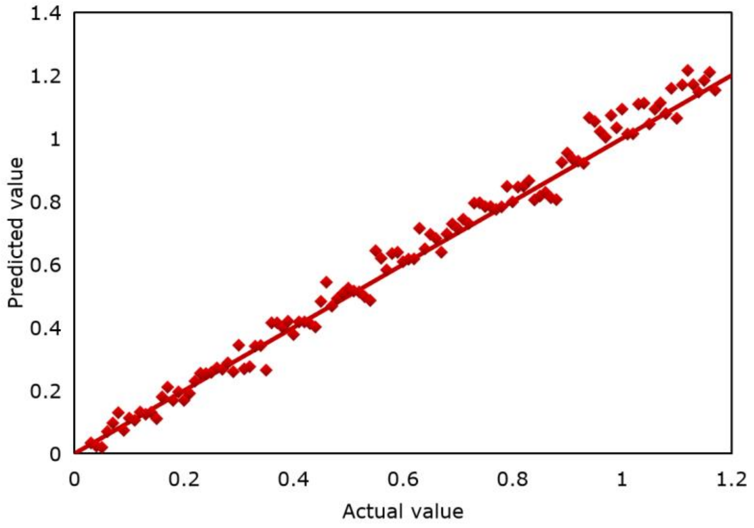

Figure 5: Accuracy test results of ACNN algorithm.

ACNN's semantic segmentation accuracy is significantly higher than traditional CNN. ACNN can continuously extract features from shallow features without reducing the resolution of feature map. In this way, the detailed information of the feature map can be well preserved, which is beneficial to capture more small-scale objects.

Figure 6 is an example of a new media video. There are 510 frames. This includes a sudden shot change at frame 280. There is one dissolution transformation at frame 42 -58. For the continuous transformation around 200 frames and 420 frames, the continuous transformation is caused by the translational movement of the lens.

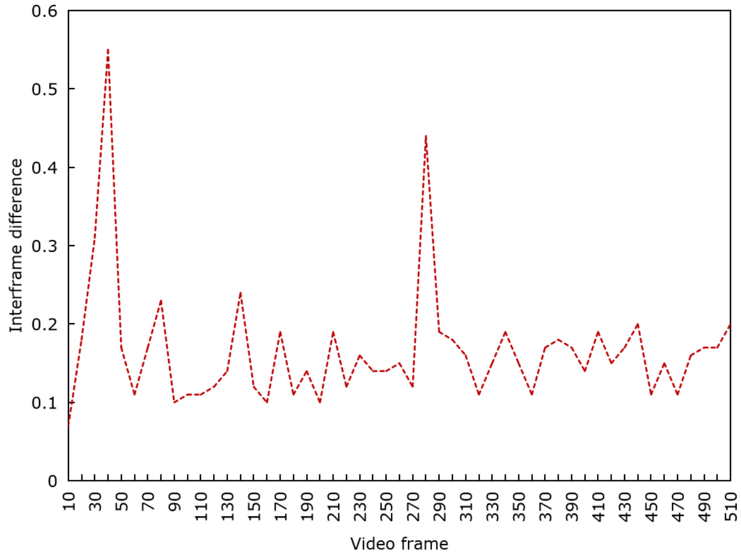

Figure 6: Video example frame difference.

By retaining more spatial information and aggregating multi-scale feature information, the model can effectively segment objects with different scales in the image, thus further improving the search efficiency of image content. The search efficiency comparison of the new media art image optimization algorithm is shown in Figure 7.

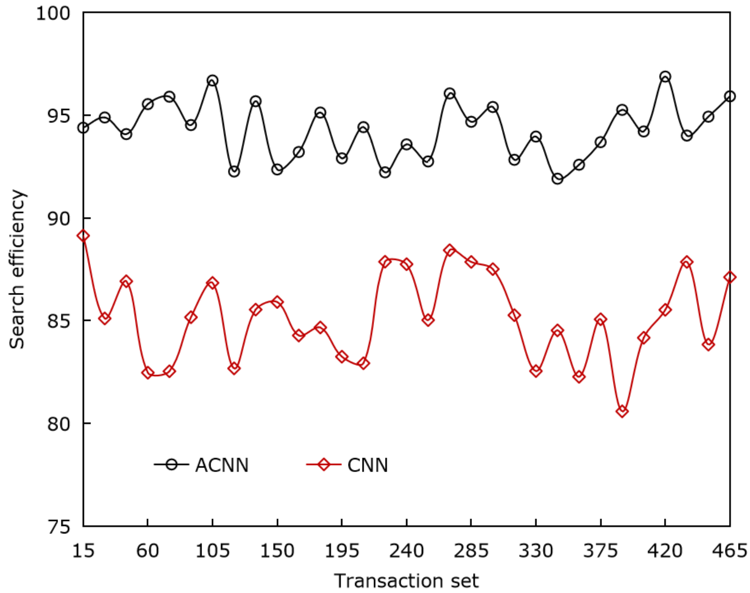

Figure 7: Comparison of search efficiency of algorithms.

According to the experimental data, the search efficiency of this algorithm is important than that of the traditional CNN model. ACNN completely preserves the computing structure of CNN by hollowing out the convolution operation and pooling operation in CNN, the comprehensive test results show that the new media art image segmentation algorithm based on ACNN algorithm has different degrees of improvement in accuracy and efficiency compared with the traditional CNN algorithm. This method obtains ideal feature recognition results of new media art images, and the segmentation accuracy is higher than other semantic segmentation methods.

6 CONCLUSIONS

New media art is obviously different from traditional art design in the design process, but pays attention to interactive characteristics. Because CAD technology has incomparable accuracy and aesthetic effect with traditional mode, its application in the field of art design has also become frequent. Image is a method of expressing creative thinking in design, and it is the carrier of conveying design theme information and ideological content. In art design, different information content will have different expression images. This article studies the optimization method of interactive design of new media art combined with CAD technology. In order to realize the real end-to-end CNN, and to directly realize the semantic segmentation result, a full ACNN is proposed. The results show that the new media art image segmentation algorithm based on ACNN algorithm has different degrees of improvement in accuracy and efficiency compared with the traditional CNN algorithm. Through this model, the fusion and reconstruction of artistic image style and photo content can be realized, in order to better enhance the interactive effect in artistic design works. By retaining more spatial information and aggregating, the model can effectively segment objects with different scales in the image, thus further improving the search efficiency of image content.

The network architecture designed in this article has the problems of large amount of parameters and calculation, so further research on network optimization can be considered, and the complexity of the network can be reduced by optimizing the network structure or model parameters, thus reducing the space and time overhead.

Huijuan Ren, https://orcid.org/0009-0005-0232-1320
Xinchun Huang, https://orcid.org/0009-0000-7520-9501

REFERENCES

[1] Benjdira, B.; Bazi, Y.; Koubaa, A.; Ouni, K.: Unsupervised domain adaptation using generative adversarial networks for semantic segmentation of aerial images, Remote Sensing, 11(11), 2019, 1369. https://doi.org/10.3390/rs11111369

[2] Chen, G.; Li, L.; Jin, W.; Zhu, J.; Shi, F.: Weighted sparse representation multi-scale transform fusion algorithm for high dynamic range imaging with a low-light dual-channel camera, Optics Express, 27(8), 2019, 10564-10579. https://doi.org/10.1364/OE.27.010564

[3] Ghosh, S.; Das, N.; Das, I.; Maulik, U.: Understanding deep learning techniques for image segmentation, ACM Computing Surveys (CSUR), 52(4), 2019, 1-35. https://doi.org/10.1145/3329784

[4] Gu, R.; Wang, G.; Song, T.; Huang, R.; Aertsen, M.; Deprest, J.; Zhang, S.: CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation, IEEE Transactions on Medical Imaging, 40(2), 2020, 699-711. https://doi.org/10.1109/TMI.2020.3035253

[5] Gupta, S.-S.; Hossain, S.; Kim, K.-D.: HDR-like image from pseudo-exposure image fusion: A genetic algorithm approach, IEEE Transactions on Consumer Electronics, 67(2), 2021, 119-128. https://doi.org/10.1109/TCE.2021.3066431

[6] Hayat, N.; Imran, M.: Detailed and enhanced multi-exposure image fusion using recursive filter, Multimedia Tools and Applications, 79(33-34), 2020, 25067-25088. https://doi.org/10.1007/s11042-020-09190-0

[7] Khened, M.; Kori, A.; Rajkumar, H.; Krishnamurthi, G.; Srinivasan, B.: A generalized deep learning framework for whole-slide image segmentation and analysis, Scientific Reports, 11(1), 2021, 11579. https://doi.org/10.48550/arXiv.2001.00258
[8] Lei, T.; Jia, X.; Liu, T.; Liu, S.; Meng, H.; Nandi, A.-K.: Adaptive morphological reconstruction for seeded image segmentation, IEEE Transactions on Image Processing, 28(11), 2019, 5510-5523. https://doi.org/10.1109/TIP.2019.2920514
[9] Ma, F.; Gao, F.; Sun, J.; Zhou, H.; Hussain, A.: Attention graph convolution network for image segmentation in big SAR imagery data, Remote Sensing, 11(21), 2019, 2586. https://doi.org/10.3390/rs11212586
[10] Merianos, I.; Mitianoudis, N.: Multiple-exposure image fusion for HDR image synthesis using learned analysis transformations, Journal of Imaging, 5(3), 2019, 32. https://doi.org/10.3390/jimaging5030032
[11] Müller, D.; Kramer, F.: MIScnn: a framework for medical image segmentation with convolutional neural networks and deep learning, BMC Medical Imaging, 21(1), 2021, 1-11. https://doi.org/10.1186/s12880-020-00543-7
[12] Niu, B.; Qu, X.; Guan, X.; Zhang, F.: Fast HDR image generation method from a single snapshot image based on frequency division multiplexing technology, Optics Express, 29(17), 2021, 27562-27572. https://doi.org/10.1364/OE.434950
[13] Qi, G.; Chang, L.; Luo, Y.; Chen, Y.; Zhu, Z.; Wang, S.: A precise multi-exposure image fusion method based on low-level features, Sensors, 20(6), 2020, 1597. https://doi.org/10.3390/s20061597
[14] Sakaridis, C.; Dai, D.; Van, G.-L.: Map-guided curriculum domain adaptation and uncertainty-aware evaluation for semantic nighttime image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(6), 2020, 3139-3153. https://doi.org/10.1109/TPAMI.2020.3045882
[15] Ulku, I.; Akagündüz, E.: A survey on deep learning-based architectures for semantic segmentation on 2d images, Applied Artificial Intelligence, 36(1), 2022, 2032924. https://doi.org/10.1080/08839514.2022.2032924
[16] Yin, J.-L.; Chen, B.-H.; Peng, Y.-T.; Hwang, H.: Automatic intermediate generation with deep reinforcement learning for robust two-exposure image fusion, IEEE Transactions on Neural Networks and Learning Systems, 33(12), 2021, 7853-7862. https://doi.org/10.1109/TNNLS.2021.3088907
[17] Zhang, L.; Wang, X.; Yang, D.; Sanford, T.; Harmon, S.; Turkbey, B.; Xu, Z.: Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation, IEEE Transactions on Medical Imaging, 39(7), 2020, 2531-2540. https://doi.org/10.1109/TMI.2020.2973595