

Research Article

Application of Optimized Convolution Neural Network Model in Mural Segmentation

Zhiqiang Chen ¹, Leelavathi Rajamanickam,¹ Xiaodong Tian ² and Jianfang Cao ^{2,3}

¹Information Technology, SEGi University, Kota Damansara 47810, Malaysia

²School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China

³Department of Computer Science and Technology, Xinzhou Normal University, Xinzhou 034000, China

Correspondence should be addressed to Zhiqiang Chen; czq_2628@163.com

Received 12 January 2022; Accepted 1 April 2022; Published 14 April 2022

Academic Editor: Soumya Ranjan Nayak

Copyright © 2022 Zhiqiang Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To address the problems of blurred target boundaries and inefficient image segmentation in ancient mural image segmentation, a multi-classification image segmentation model MC-DM (Multi-class DeeplabV3+ MobileNetV2) that fuses lightweight convolutional neural networks is proposed. The model combines the DeeplabV3+ structure and MobileNetV2 network and adopts the unique spatial pyramid structure of DeeplabV3+ to process convolutional features for multi-scale fusion, which reduces the loss of detail in the mural segmentation images. Firstly, the features calculated at any resolution in MobileNetV2 network are extracted by hole convolution, the input step is expressed as the ratio between the input image resolution and the final resolution, and the density of encoder features is controlled according to the budget of computing resources. Then, the spatial pyramid pool structure is used to fuse the previously calculated features at multiple scales to enrich the semantic information of the feature image. Finally, the same convolution network is used to reduce the number of channels and filter the density feature map. The filtered features are fused with the features after multi-scale fusion again to obtain the final output. In total, 1000 scanned images of murals were adopted as datasets for testing under the JetBrains PyCharm Community Edition 2019 environment. The obtained experimental results indicate that MC-DM improves the training accuracy by 1 percentage point compared with the conventional SegNet-based image segmentation model, and by 2 percentage points compared with the PspNet network-based image segmentation model. The PSNR (peak signal-to-noise ratio) of the MC-DM model is improved by 3–8 dB on average compared with the experimental model. This confirms the effectiveness of the model in mural segmentation and provides a novel method for ancient mural image segmentation.

1. Introduction

Ancient murals are mediums of Chinese culture and have significant historical value; however, under natural and human impact, the ancient murals of the distant past have been exposed to various degrees of damage, while their content have been severely damaged. Hence, image restoration in murals has become one of the most difficult problems faced by cultural workers and historical researchers in the course of analyzing ancient murals. Mural segmentation is the first step of image analysis and plays a very important role in image engineering. Image feature extraction, target recognition and target detection all depend on the quality of image segmentation in the later image

process stage. Similarly, as a key step of mural digital protection, mural segmentation is the basis of mural classification and restoration. The segmentation results directly affect the process of cultural relics protection. Therefore, the research on mural segmentation methods has attracted more and more attention.

Deep learning, a learning method based on artificial neural networks that imitates the human brain to process and interpret data, is a new field in machine learning research and is widely used in several fields such as image and sound processing. Deep learning can combine neural networks with probabilistic models to improve the inference ability of image models; hence, in the field of image segmentation, various image segmentation models based on

deep learning have been proposed to effectively solve series of problems such as blurred image edge segmentation and missing information of segmented images in conventional segmentation methods. Based on the above, this paper proposed a new DeeplabV3+ model by improving the deep learning model and applied it to the segmentation of ancient murals.

2. Related Works

About the application of deep learning model in the field of image segmentation, researchers adopted fully convolutional networks (FCN) [1] or improved FCN networks for image segmentation initially, which featured a fully connected layer of convolutional neural networks (CNN) [2] replaced with a convolutional layer to adapt to arbitrary size inputs and output low-resolution segmented images. However, this method has significant limitations: the edge segmentation of FCN is poor, and the contour of the segmented image is blurred. Chen et al. [3] proposed the DeeplabV1 model in 2015 to address this problem, which utilizes a fully-connected conditional random field (CRF) to optimize boundary segmentation and effectively solve the edge contour segmentation problem inherent in FCN. The DeeplabV3+ [4] model is a modification of the previous generation DeeplabV3 model; it is a novel and improved scheme that can help researchers refine segmentation results and fares better in the delineation of object boundaries. In 2019, Ren et al. [5] combined the DeeplabV3+ model with the super pixel segmentation algorithm, simple linear iterative cluster (SLIC), and experimentally demonstrated that DeeplabV3+ has a better image detail restoration ability than FCN and SegNet segmentation [6] models. Image segmentation technology based on deep learning has been developing. Especially since the 2020 COVID-19, image segmentation technology has developed rapidly in medicine [7–10], which provides a new way of thinking for ancient murals.

In ancient mural image segmentation, conventional segmentation methods are mostly used, and these segmentation models are not universally applicable. Conventional mural segmentation methods apply various approaches, and one of them involves using the fuzzy C-mean (FCM) [11–13]. This objective-based fuzzy clustering algorithm is widely used, and its algorithm theory is mature; however, if this algorithm is employed in the mural segmentation field, it will be affected by sample imbalance. When the sample capacity of different classes is not consistent, it will cause a certain class of segmentation samples to encounter difficulties in approaching the target sample, which triggers poor segmentation results. The second approach is the mean drift algorithm-mean shift [14–17], which is essentially an estimation algorithm for kernel density; however, this algorithm runs slowly, and is only applicable to feature data point sets for which standard features have been established in mural segmentation; in addition, it is prone to the presence of images outside the target or missing parts of the target, and has a limited effect when performing batch segmentation. The third

conventional mural segmentation algorithm, graph cuts [18–20], adopts a graph form to solve the energy function and assign corresponding weights to the edges of the graph and transform the energy function into an S/T graph for complete image segmentation. However, this method exhibits a poor segmentation effect when handling noise or occlusion, and it requires manual labeling of a number of front and back view pixel points, which has a series of problems such as manual intervention [21].

Based on the efficiency of deep learning neural networks, this study proposes a multi-class lightweight network segmentation model (Multi-class DeeplabV3+ MobileNetV2, MC-DM) that combines the lightweight convolutional neural network MobileNetV2 [22] with the DeeplabV3+ model. The model uses the DeeplabV3+ structure to collect multi-scale information of the image, effectively circumventing the missing semantic information of the image; in addition, it adopts the MobileNetV2 convolutional neural network to extract features, improve the efficiency of mural segmentation, and reduce the influence of hardware conditions on the segmentation effect. Experiments indicate that the method has different degrees of segmentation accuracy and efficiency in the mural image segmentation process and exhibits optimal robustness in terms of image segmentation edge continuity.

3. Materials and Methods

The improved MC-DM model involves lightweight neural network MobileNetV2, DeeplabV3+ model, ASPP structure and other related concepts. Therefore, we divide two parts to introduce: Relevant theories and Mural segmentation model MC-DM. Relevant theories focuses on the working principle of relevant network and model structure; Mural segmentation model MC-DM introduces the improvements and excellent characteristics of the proposed model.

3.1. Relevant Theories. In this part, we discuss the network structure, convolution mode of MobileNetV2 network and the working principle of DeeplabV3+ model respectively.

3.1.1. MobileNetV2. The MobileNetV2 convolutional neural network is proposed to solve the problems of large convolutional neural networks and insufficient hardware training that emerge during the training of image models. It is an important approach to addressing the hardware memory limitation of deep learning models deployed in mobile devices [23]. It is another important invention after SqueezeNet [24], ShuffleNet [25], Xception [26], and other lightweight neural convolutional networks. The core part of the network is depthwise separable convolution, and its operation comprises two parts: depthwise (DW) and pointwise (PW) convolution. With a 3×3 convolution kernel and a large number of channels, DW separable convolution can reduce the computational effort by approximately 9 times less than normal convolution.

Based on the first-generation lightweight network MobileNetV1, the MobileNetV2 network introduces the

concepts of inverted residuals and linear bottlenecks; this limits the feature extraction in terms of the number of input channels because DW convolution does not change the number of channels. These two parts adopt the low dimensional compression as input, expand it to a high dimension, and then filter it using lightweight deep convolution; subsequently, the resulting features are represented by linear convolution projection into low dimension. The net structure of the MobileNetV2 is presented in Table 1.

In Table 1, t , c , n , and s denote the dilation factor, number of output channels, number of repetitions of the convolution layer, and step size, respectively. The first layer of each sequence has a step size, all other layers adopt a step size of 1, and all spatial convolutions employ a 3×3 convolution kernel. Each bottleneck contains three parts: dilation, convolution, and compression, with each row describing one or more sequences, repeated n times, and all layers in the same sequence having the same number of output channels. The MobileNetV2 network facilitates a significant reduction in the memory footprint problem required during inference by not fully specifying the intermediate tensor, and its application to mural segmentation can reduce the need for main memory accesses in most embedded hardware designs.

3.1.2. Conventional DeeplabV3+ Model. The DeeplabV3+ model is an improvement of the DeeplabV3 model with the residual neural network (ResNet) network as the underlying network, which also encodes an encoder-decoder structure to obtain clear object boundaries by recovering spatial information to optimize boundary segmentation. The ResNet network or Xception network is used for the feature extraction of the input image, after which, the image special is fused via atrous spatial pyramid pooling (ASPP) to prevent information loss. In the DeeplabV3+ model, the DeeplabV3 model is adopted as the encoder part with an external simple and effective decoder module to obtain clear results.

Null convolution with multiple null rates (rates) is employed in Deeplabv3+ to efficiently extract contextual information in parallel, and this structure adopts the ASPP model to provide multi-scale information. The structure of this model is illustrated in Figure 1.

The ASPP module comprises a 1×1 convolution and three 3×3 null convolutions with sampling rates of 6, 12, and 18, respectively. In the Deeplabv3+ model, the input image is divided into two parts after passing through the backbone deep neural convolutional network, with one part going into the decoder and the other part going into the parallel null convolutional structure i.e., the ASPP model. Separate feature extraction is performed with different rates of void convolution, which is then merged, followed by 1×1 convolution, for which feature compression is performed. Then, the compressed feature map is upsampled four times via bilinear interpolation, to pass into the decoder.

3.2. Mural Segmentation Model MC-DM. In this section, we focus on the improvement of MC-DM model and the working principle of each part of the model.

TABLE 1: Structure of MobileNetV2 network.

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	—	32	1	2
$112^2 \times 32$	Bottleneck	1	16	1	1
$112^2 \times 16$	Bottleneck	6	24	2	2
$56^2 \times 24$	Bottleneck	6	32	3	2
$28^2 \times 32$	Bottleneck	6	64	4	2
$14^2 \times 64$	Bottleneck	6	96	3	1
$14^2 \times 96$	Bottleneck	6	160	3	2
$7^2 \times 160$	Bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1×1	—	1280	1	1
$7^2 \times 1280$	Avgpool 7×7	—	—	1	—
$1 \times 1 \times 1280$	conv2d 1×1	—	k	—	—

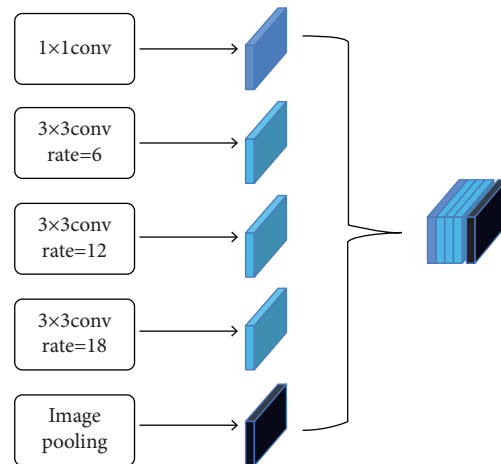


FIGURE 1: ASPP model structure diagram.

3.2.1. DeeplabV3+ MC-DM Incorporating MobileNetV2. The DeeplabV3+ underlying network is highly adaptive. To achieve segmentation accuracy, researchers have incorporated ResNet. Although such models have high classification accuracy, their model depth keeps deepening, which triggers an increase in model complexity. Complex segmentation models are constrained by hardware memory and are demanding for mobile or embedded devices, which cannot satisfy the segmentation requirements of low latency and high response rate in specific scenarios. To address this problem, a segmentation model that combines the lightweight neural network MobileNetV2 with the segmentation model DeeplabV3+ is proposed. The encoder module in the model is employed to reduce feature loss and capture higher-level semantic information, while the decoder module is used to extract details and recover spatial information. The model decomposes the convolution into two independent layer factors to replace the full convolution operator, performs light filtering by applying a single convolution filter to each input channel, and later constructs new features via the linear combination of the input channels. The changes to the convolutional network optimize the performance of the DeeplabV3+ decoder module in recovering detailed object boundaries.

With the same dataset, the MC-DM model adopts a network that has a significant advantage in segmentation

efficiency over convolutional networks such as ResNet and Xception. The most significant difference between this model and the conventional DeeplabV3+ is that instead of using standard convolution to extract features, the DW convolution that can perform feature extraction in high dimensions is adopted. The advantage of this method is that it makes the MC-DM model substantially less computationally intensive than the conventional DeeplabV3+ model, which can be applied in the mural segmentation field to satisfy the efficient requirements of mural segmentation while ensuring accuracy. The improved model is illustrated in Figure 2.

The first improvement of the model is the combination of hole convolution structure and deep separable network structure. In Figure 2, structure A represents the null convolution, which extracts features computed at arbitrary resolution from the MobileNetV2 network, expresses the input step size as the ratio obtained from the input image resolution to the final resolution, and controls the density of encoder features based on the budget of computational resources, to control the budget for encoder computational resources. For the semantic segmentation task, an output with a step size of 16 is used for more intensive feature extraction after discarding the span in the last one or two blocks. This approach is taken because when the decoder output step size is 8, the segmentation performance is improved relative to the output step size of 16, and although the performance is improved, it increases the computational complexity. Therefore, in the MC-DM, the output step size used for the encoder module is 16, which has the advantage of balancing segmentation accuracy and speed.

The second improvement of the model is combining the spatial pyramid pool with MobilNetV2, as shown in structure B. The structure uses hole convolution with different hole rates to fuse the features calculated by MobilNetV2 at multiple scales, which enriches semantic information and effectively balances accuracy and running time.

The third improvement of the model is using the same convolution network to reduce the number of channels and modify the output step of the model, as shown in structure C. The use of the same convolution network solves the problem of training difficulty caused by a large number of channels in low-level features. Secondly, we modify the output step setting and set its value to 4, which can make appropriate trade-offs for density feature mapping and simplify the Decoder module under the condition of limited GPU resources, so as to improve the image segmentation efficiency of the model.

3.2.2. Description of the Algorithm. The workflow of the MC-DM segmentation model is presented in the following steps.

Step 1: Input the mural image of the fixed size and resolution into the segmentation model.

Step 2: Perform the feature extraction of the image using improved depth separable network and retain the mural image detail information using null convolution.

Step 3: Shunt low-level features into the ASPP and Decoder structures, respectively, to retain image feature information to a great extent.

Step 4: Multi-scale-fuse the feature information passing through the ASPP structure via 1×1 convolution and feed the fusion result into the decoder structure; the low-level features that initially enter the decoder structure are refined by different convolution layers to refine the features.

Step 5: Upsample the encoder output feature map via bilinear interpolation and maintain the same size as that of the feature map after the feature refinement in decoder. The sampled results are fused with the refined results and features to obtain a more feature-rich mural image.

Step 6: Upsample the feature fusion image again to obtain a segmented image with the same parameters as the input image, and then complete the segmentation process.

4. Analysis of Experimental Results

4.1. Experimental Environment and Data Sources. The personal computer environment for the experiments is Windows 10 with Intel Core i7-9750H CPU, NVIDIA GeForce 1660Ti GPU, and 8G RAM. The TensorFlow deep learning framework was used to train and test the semantic segmentation model in the text.

The dataset of DeeplabV3+ employed a single-channel annotated map, and the experimental images were obtained from the scanned images of the album “The Complete Collection of Dunhuang Murals in China,” while the image annotation of the scanned images was developed into a dataset by the graphic user interface annotation software, LabelMe. The sample dataset graph is presented in Figure 3.

Figure 3(a) represents the scanned image, based on which the edges of the scanned image are labeled point by point using floating points and the labeled points are connected to form the result presented in Figure 3(b). Subsequently, based on the original and annotated images, a single-channel grayscale image was trained and merged with the scanned image to form the dataset. The dataset contains 1000 images divided into five categories: animals, houses, people, auspicious clouds, and Buddha images, with 200 images in each category. The images were pre-processed using the letterbox function to prevent missing frames in the images during the training process. To reduce the occurrence of overfitting triggered by few images, experiments were performed to enhance the obtained images. The enhancement was carried out by changing the color of the image, increasing the noise, and changing the brightness. Figure 4 presents the image obtained from the data enhancement.

The original image is presented in column (a) of Figure 4, while the last four columns depict the enhanced images. The obtained results need to be tested several times owing to the stochastic nature of the enhancement with functions. In the experimental phase, 90% of the dataset was adopted for training and 10% for prediction. The experiments were

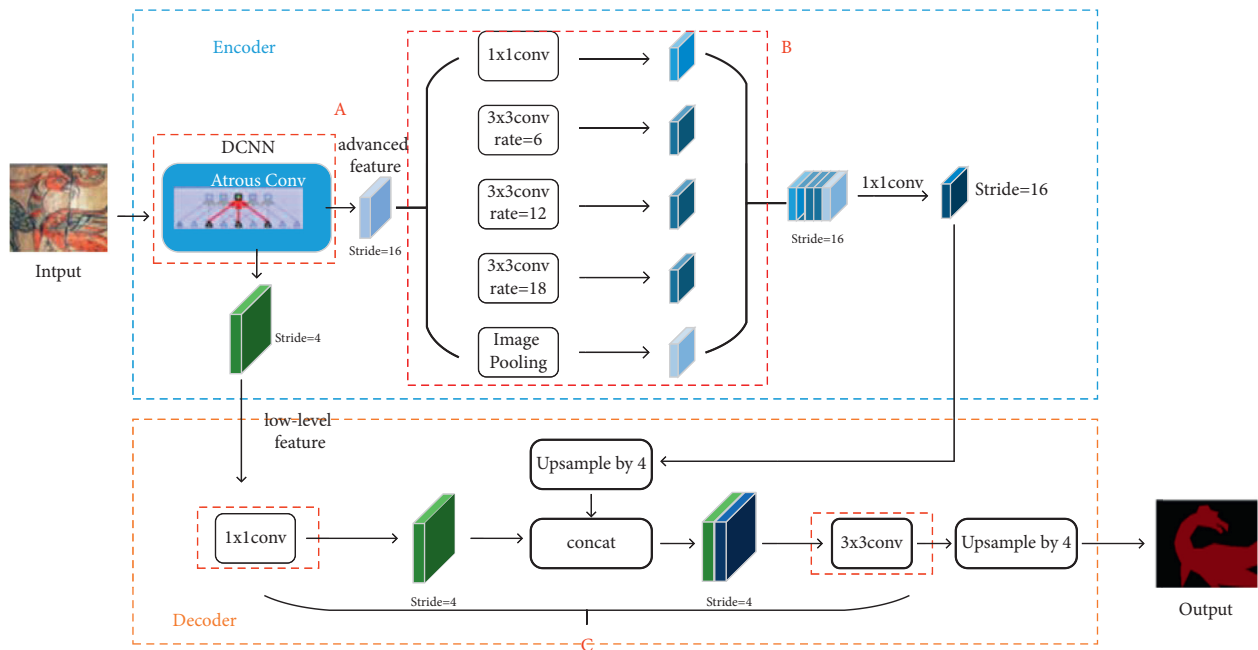


FIGURE 2: MC-DM diagram.

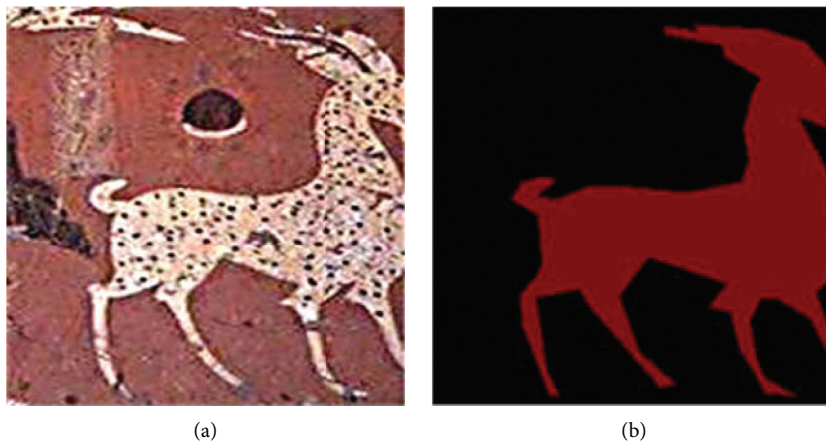


FIGURE 3: Sample plot of DeeplabV3+ dataset. (a) Scanned image of an ancient mural. (b) Label image of an ancient mural.

limited to the accuracy of the test set, and when the loss value val_loss of the test set did not decrease twice in a row, the learning rate was reduced to continue the training. Training was cut off when the loss value stabilized, and the obtained data were saved every 30 generations. The variation of the splitting accuracy is presented in Figure 5.

To improve the experimental training accuracy, the first 10 generations of test set loss values took a wide range, which triggered large fluctuations in the experimental test set training accuracy. After 10 generations, the overall accuracy of the experiments and the test set training accuracy gradually increased and stabilized at the 40th generation, while the learning rate reached the optimum.

4.2. Comparative Experiments. Three different image segmentation models were designed for comparisons with the

models presented in this study, based on the homemade dataset. First, the MobileNetV2 network was combined with the models in [27, 28] as comparison Models 1 and 2, respectively. The model in [29] was adopted as comparison Model 3. All three models were altered to ensure that one part of the combined model remained unchanged and optimally comparable. Four images of different types were selected from the dataset for segmentation, and the segmentation results were labeled at pixel level to obtain the visual comparison effect, and the obtained results are presented in Figure 6.

In Figure 6, column (a) shows the image to be segmented, columns (b), (c), and (d) illustrates the segmented images of the original image under Models 1, 2, and 3, respectively, and column (e) presents the segmentation results of the MC-DM model. In Model 1, owing to the use of continuous downsampling resulting in a large amount of



FIGURE 4: Data enhancement image. (a) Original image. (b) Upside down. (c) Flipped. (d) Noise(a). (e) Darken.

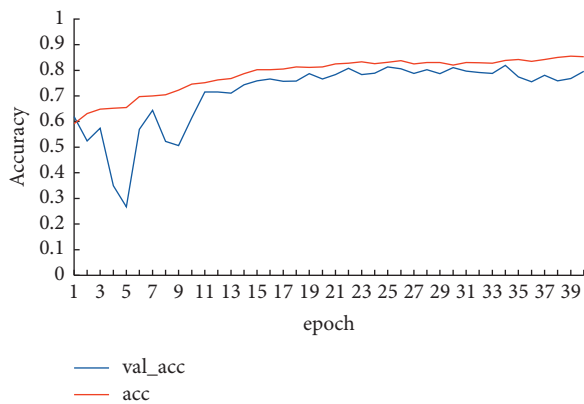


FIGURE 5: Accuracy variation graph.

spatial information in the input image overlapping each one-pixel on the output feature map, multiple image spatial information with lossy boundary information is not conducive for image segmentation. Model 2 first performs multi-scale pooling of the input feature information, after which the pooling results are upsampled, and then stitched. The advantage of this approach is that the information of different sensory fields can be utilized to enrich the image content; however, it easily triggers a situation where severe loss of single-category image information occurs, and the segmented edges do not match the real edges, as illustrated in Figure 6(c). Model 3 combines the DeeplabV3+ model and Xception network, which ignorantly increases the number of convolutional network parameters, thereby increasing the difficulty of image training; in addition, the

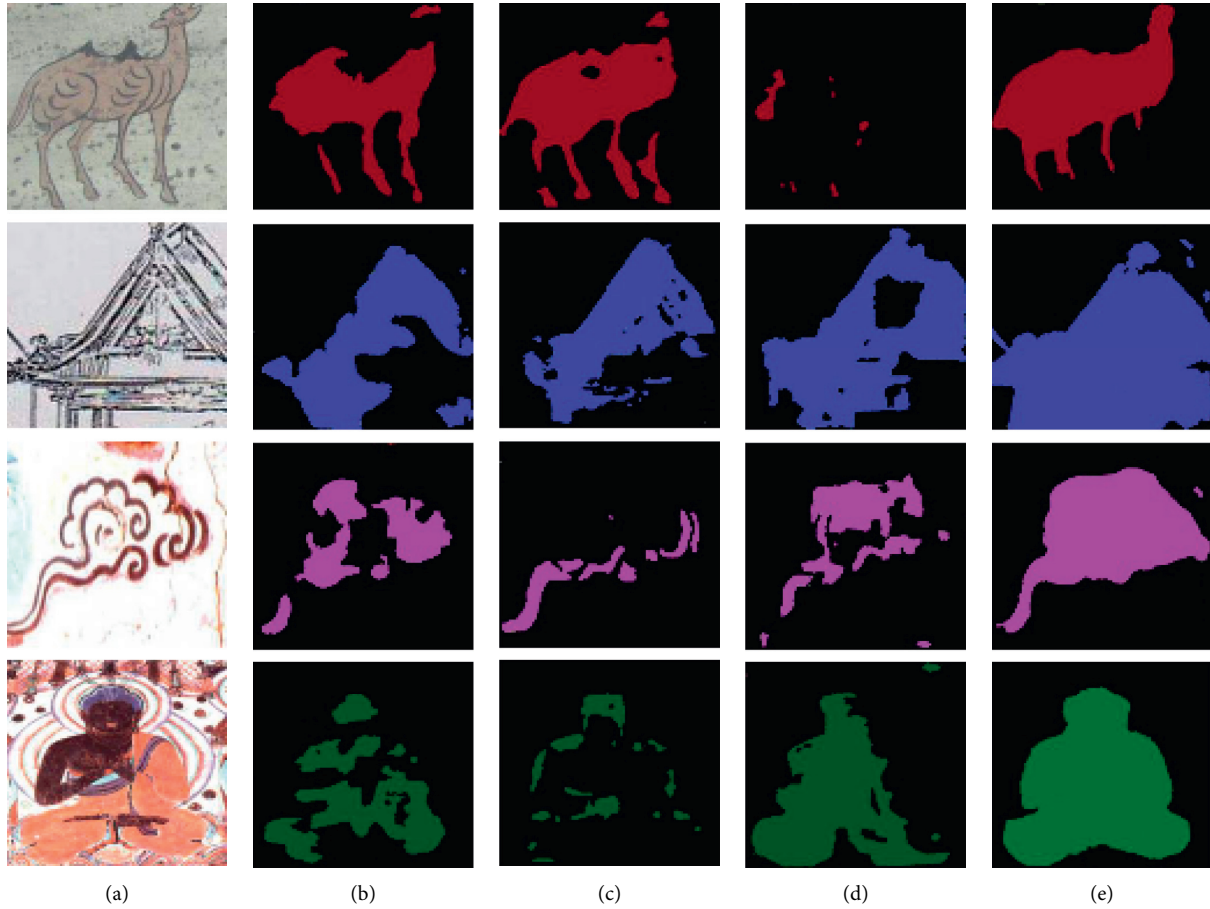


FIGURE 6: Comparison of segmentation effects. (a) Sample image. (b) Model 1. (c) Model 2. (d) Model 3. (e) MC-DM.

image segmentation results are influenced by the hardware equipment, and the loss of details in the center of the segmented image is severe. The MobileNetV2 network used by the MC-DM segmentation model reduces the number of networks. Furthermore, increasing the decoder structure extracts the image details, and the segmentation effect is the best among the four models.

The peak signal-to-noise ratio (PSNR) is adopted as an objective indicator, and the magnitude of the value represents the frame loss rate of the segmented image. Accordingly, a higher value represents a better image segmentation effect. The results of the PSNR values for four randomly selected samples are presented in Table 2.

In Sample 1, the sample image lines are simple, the four segmentation models have similar effects, and the MC-DM segments the image with the highest PSNR value, which is 1 dB higher than the comparison model. Samples 2 and 3 have relatively complex image contours, and partial fusion exists between the target and background, while MC-DM segments the image with a significant increase in the PSNR value, which is 5 dB higher than the comparison model on average. Sample 4 has an image with a complex structure and more background information, which exerts a more significant impact on the segmentation results of the image. MC-DM performs well in the segmentation results of this sample, and the PSNR values are improved by 10 dB on

TABLE 2: PSNR (dB) comparison.

Sample	Model 1	Model 2	Model 3	MC-DM
1	25.86	26.39	21.42	26.66
2	16.08	15.91	17.85	20.00
3	16.67	14.61	15.98	21.98
4	20.85	18.62	21.98	30.73

average, compared with the comparison model; in addition, the experiment verifies the feasibility of this model in mural segmentation. The training accuracies of the four models are presented in Table 3.

Model 1 adopts deconvolution and up-pooling, which can only barely recognize the image shape, and its segmentation results are coarse. Model 2 has more missing details in the center of the image, although features of different sizes are obtained by multi-scale pooling. Model 3 improves the underlying network of the model; in addition, using depth-separable convolution, it optimizes the feature extraction method in the mural image segmentation process, but its segmentation results are poor for a single-category image. The improved model, MC-DM, is the most efficient in the mural segmentation process and it addresses the problem of missing details in Model 2. Compared with Model 1, MC-DM-segmented image edges are completely preserved, and the loss of image information is not

TABLE 3: Training accuracy table.

Segmented model	Accuracy/%
Model 1	84.07
Model 2	83.59
Model 3	84.88
MC-DM	85.30

significant. MC-DM model exhibits better applicability than Model 3 and does not exhibit the phenomenon of large differences in segmentation results due to different image types. Hence, it can be inferred from the two experimental parameters of PSNR and training accuracy that the segmentation effect of MC-DM is better than that of the other three models, and the model segmentation contour tends to the ideal contour without causing a large number of missing details.

5. Conclusions

Ancient Chinese murals are an important witness of Chinese civilization and an inseparable part of the development of the history of world civilization. Due to the long history, murals are negatively affected by many factors such as environment and man-made, and there are many problems such as image deformity, falling off and cracks. How to effectively preserve these precious cultural relics is the top priority at present. In this paper, the deep learning model is integrated into mural image segmentation, and the powerful learning ability of neural network is used to improve the problems of traditional segmentation methods, such as fuzzy image edge segmentation, which is a new exploration in ancient mural image processing. The main contributions of this paper are reflected in the following two aspects: (1) MC-DM model is proposed. The model uses hole convolution and lightweight neural network to extract mural image features, adjust the output step of Decoder structure, and balance the accuracy and speed of network segmentation. The same convolution network is used to reduce the number of channels. The density feature mapping is properly selected to reduce the difficulty of model training in the case of limited GPU resources. (2) The proposed MC-DM model is applied to the segmentation of ancient murals to solve the problems of unclear segmentation target boundary and low segmentation efficiency of traditional mural segmentation models. Based on the idea of deep learning, this paper carries out image segmentation of ancient murals, makes a systematic research on feature extraction and feature fusion on the basis of the original research, improves the ability of image feature restoration, effectively interprets the mural image meaning, and provides a new idea for the research of digital protection of ancient cultural relics.

However, there are still some problems in the experimental process, such as small data scale, lack of feature information, poor effect of multi sharp point image segmentation and so on. The model proposed in this paper still needs to be improved to meet the changing practical requirements. Therefore, the future work will be carried out from the following two aspects: (1) In the experimental stage,

because DeeplabV3+, like other models in the Deeplab family, requires a specific dataset, the samples need to be manually labeled in the early stage, which is a huge workload. This problem can be solved by continuing to collect high-quality mural images with high quality and rich image information, and constantly expanding the number of images in the data set. The continuous improvement of the data set will make the training of the model more sufficient and better, which could avoid the problems of over fitting and under fitting caused by the lack of data. (2) The problem of blurred edges emerges in the multi-category image segmentation due to the geometric reduction of the experimentally encoded output feature maps relative to the input images. This is also a problem that needs to be further addressed in the future segmentation of ancient wall paintings.

Data Availability

All data used for analysis in this study are included within the article.

Conflicts of Interest

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the Project of Key Basic Research in Humanities and Social Sciences of Shanxi Colleges and Universities (20190130).

References

- [1] Z. Wu, C. Shen, and A. van den Hengel, "Wider or deeper: revisiting the resnet model for visual recognition," *Pattern Recognition*, vol. 90, no. 6, pp. 119–133, 2019.
- [2] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Boston, MA, USA, June 2015.
- [3] L.-C. Chen, Y.-K. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 801–818, Munich, Germany, September 2018.
- [4] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.
- [5] F.-L. Ren, X. He, Z.-H. Wei, L. You, and L. I. Mu-Yu, "Semantic segmentation based on DeepLabV3+ and superpixel optimization," *Optics and Precision Engineering*, vol. 27, no. 12, pp. 2722–2729, 2019.
- [6] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [7] S. Dash, M. R. Senapati, P. K. Sahu, and P. S. R. Chowdhary, "Illumination normalized based technique for retinal blood

- vessel segmentation,” *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, 2020.
- [8] A. Ranganath, M. R. Senapati, and P. K. Sahu, “Classification of textures using pixel range calculation method,” in *Proceedings of the 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, Greater Noida, India, February 2021.
- [9] A. Ranganath, P. K. Sahu, and M. R. Senapati, “Detection of COVID from chest X-ray images using pivot distribution count method,” in *Proceedings of the 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 373–378, Noida, India, August 2021.
- [10] A. Ranganath, P. K. Sahu, and M. R. Senapati, “A novel approach for detection of coronavirus disease from computed tomography scan images using the pivot distribution count method,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 10, no. 2, pp. 145–156, 2021.
- [11] R. Lan and Y. Lin, “Suppressed nonlocal space Intuitionistic Fuzzy C-Means Image segmentation algorithm,” *Journal of electronics and information*, vol. 41, no. 6, pp. 1472–1479, 2019.
- [12] J.-J. Sun and Y. Xu, “Underdetermined mixed matrix estimation based on weighted improved fuzzy C-means clustering,” *Computer Applications*, vol. 40, no. 6, pp. 1769–1773, 2020.
- [13] S. Mishra, P. Sahu, and M. R. Senapati, “MASCA-PSO based LLRBFNN model and improved fast and robust FCM algorithm for detection and classification of brain tumor from MR image,” *Evolutionary Intelligence*, vol. 12, no. 4, pp. 647–663, 2019.
- [14] S.-N. Zhao and W.-J. Wang, “Image segmentation algorithm based on SVM and fast mean shift,” *Journal of Small Microcomputer Systems*, vol. 38, no. 7, pp. 1614–1618, 2017.
- [15] I. A. Iswanto, T. W. Choa, and B. Li, “Object tracking based on meanshift and particle-kalman filter algorithm with multi features,” *Procedia Computer Science*, vol. 157, pp. 521–529, 2019.
- [16] M. Lu and Y. Xu, “Overview of target tracking algorithms,” *Acta Automatica Sinica*, vol. 45, no. 7, pp. 1244–1260, 2019.
- [17] S. Das, A. Patra, S. Mishra, and M. R. Senapati, “A self-adaptive fuzzy-based optimised functional link artificial neural network model for financial time series prediction,” *International Journal of Business Forecasting and Marketing Intelligence*, vol. 2, no. 1, p. 55, 2015.
- [18] B. Martin, D. Paulusma, and E. J. van Leeuwen, “Disconnected cuts in claw-free graphs,” *Journal of Computer and System Sciences*, vol. 113, pp. 60–75, 2020.
- [19] V. R. Balaji, S. T. Suganthi, R. Rajadevi, V. Krishna Kumar, B. Saravana Balaji, and S. Pandiyan, “Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes Classifier,” *Measurement*, vol. 163, Article ID 107922, 2020.
- [20] A. Ranganath, M. R. Senapati, and P. K. Sahu, “Estimating the fractal dimension of images using pixel range calculation technique,” *The Visual Computer*, vol. 37, no. 4, pp. 635–650, 2021.
- [21] M. Antonello, S. Chiesurin, and S. Ghidoni, “Enhancing semantic segmentation with detection priors and iterated graph cuts for robotics,” *Engineering Applications of Artificial Intelligence*, vol. 90, Article ID 103467, 2020.
- [22] M. Sandler, A. Howard, M.-L. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4510–4520, Salt Lake City, UT, USA, June 2018.
- [23] Q.-H. Li, C.-P. Li, J. Zhang, H. Chen, and S.-Q. Wang, “Survey of compressed deep neural network,” *Computer Science*, vol. 46, no. 9, pp. 1–14, 2019.
- [24] K. Nan, S. Liu, J. Du, and H. Liu, “Deep model compression for mobile platforms: a survey,” *Tsinghua Science and Technology*, vol. 24, no. 6, pp. 677–693, 2019.
- [25] F. Daghero, D.-J. Pagliari, and M. Poncino, “Energy-efficient deep learning inference on edge devices,” *Advances in Computers*, vol. 122, pp. 247–301, 2020.
- [26] J.-J. Jiang, Y.-F. Xiong, and X. Xia, “A manual inspection of Defects4J bugs and its implications for automatic program repair,” *Science China (Information Sciences)*, vol. 62, no. 10, pp. 31–46, 2019.
- [27] S.-Q. Luo, Z.-C. Zhang, and Q. Yue, “Image semantic segmentation based on improved SEGNET model,” *Computer Engineering*, vol. 47, no. 4, pp. 256–261, 2021.
- [28] J. Zhou, M. Hao, D. Zhang, P. Zou, and W. Zhang, “Fusion PSPnet image segmentation based method for multi-focus image fusion,” *IEEE Photonics Journal*, vol. 11, no. 6, pp. 1–12, 2019.
- [29] X. Tian, L. Wang, and Q. Ding, “Image semantic segmentation based on deep learning,” *Journal of Software*, vol. 30, no. 2, pp. 40–46, 2019.