

Article

UR-Net: An Optimized U-Net for Color Painting Segmentation

Zhen Liu ¹, Shuo Fan ^{2,*}, Silu Liu ³ and Li Liu ²¹ School of Communication, Qufu Normal University, Rizhao 276800, China; zhen@whu.edu.cn² School of Physics and Electronic Information, Yunnan Normal University, Kunming 650500, China; liuli@ynnu.edu.cn³ College of Engineering, Qufu Normal University, Rizhao 276800, China

* Correspondence: 13581064038@163.com

Abstract: The pigments of cultural color paintings have faded with the passage of time. Color segmentations are essential for digital color reconstruction, but the complexity of color paintings makes it challenging to achieve high-precision segmentation using previous methods. To address the challenges of color painting segmentation, an optimized strategy based on U-Net is proposed in this paper. The residual blocks of a residual network (ResNet) are added to the original U-Net architecture, and a UR-Net is constructed for the semantic segmentation of color paintings. The following steps are taken. First, datasets of color paintings are obtained as training and test samples and are labeled with the two following pixel colors: earth red and ultramarine blue. Second, residual blocks are improved and added to fit the U-Net architecture. Then, a UR-Net is constructed and trained using the samples to obtain the semantic segmentation model. Finally, the effectiveness of the trained UR-Net model for segmenting the test samples is evaluated, and it is compared with the K-means clustering algorithm, ResNet, and U-Net. Data from several studies suggest that the segmentation accuracy of the UR-Net model is higher than that of other methods for the color segmentation of painted images, and the IoUs of the segmented earth red and ultramarine blue pixels are 0.9346 and 0.9259, respectively, achieving the desired results. The proposed UR-Net model provides theoretical and methodological support for further in-depth research on color recognition and segmentation of cultural color paintings.

Keywords: color painting; color recognition and segmentation; residual blocks; U-Net; semantic segmentation



Citation: Liu, Z.; Fan, S.; Liu, S.; Liu, L. UR-Net: An Optimized U-Net for Color Painting Segmentation. *Appl. Sci.* **2024**, *14*, 10005. <https://doi.org/10.3390/app142110005>

Academic Editor: Jae-Hong Choi

Received: 22 September 2024

Revised: 23 October 2024

Accepted: 30 October 2024

Published: 1 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cultural heritage has long been a subject of great interest in a wide range of fields. China has nurtured countless pieces of cultural heritage throughout its rich history, among which color painting is one of the brightest pearls. There are many examples of color paintings with visual properties that have changed with the passage of time [1]. Examples include the darkening and fading of pigments caused by the natural environment and human activities [2]. Therefore, one of the most significant challenges is reproducing the original appearance of color paintings. To date, cultural color paintings have been reconstructed by various methods, and a visual reinforcement approach has been employed using pigments to add color directly to the faded area [3]. Such approaches however, are potentially harmful, can cause secondary damage, and go against conservation ethics. Fortunately, digital color reconstruction can overcome the significant drawbacks of visual reinforcement and precisely reconstruct the original color without damaging cultural color paintings [4].

The high-precision segmentation of painted images is one of the bases of digital color reconstruction, which can distinguish complex pigments [5]. By this method, we can enable the determination of the concentration maps for each color used in a color painting by mapping the pixels of each color individually so that digital color reconstruction can be performed more efficiently and accurately. However, unique line patterns and

color combinations make painted images complex, which made it difficult for previous segmentation methods to achieve the desired accuracy and precise reconstruction of faded pigments. Therefore, we must develop a model that recognizes colors and performs high-precision segmentation.

The results from earlier studies demonstrate that the K-means clustering algorithm [6] is an integral part of unsupervised learning and is a previous method for image segmentation. Docquier et al. [7] designed a semiautomatic segmentation method based on the K-means clustering algorithm to measure the volume of a bone cyst (simple or aneurysmal) from magnetic resonance imaging (MRI). Due to the complexity of color paintings, the K-means clustering algorithm has difficulty achieving color consistency when performing global clustering on the input image, which may lead to unsatisfactory segmentation accuracy [8]. Semantic segmentation models built based on convolutional neural networks (CNNs) have been extensively researched in the field of image recognition and segmentation [9]. He et al. [10] proposed ResNet for semantic segmentation based on a CNN. This simplified the training of networks that are substantially deeper than a regular CNN. Yang et al. [11] used a lightweight network (ResNet-G18) that integrated ResNet-18 and a ghost bottleneck to develop a classification model for sound and bruised nectarines based on polarization image data. ResNet can solve degradation problems of feature extraction with an increasing number of network layers by residual blocks. However, ResNet requires many datasets in the feature extraction process to train a more accurate semantic segmentation model, and the number of samples for color painting is very sparse, which makes it difficult to achieve color consistency in segmentation.

The past decade has seen the rapid development of deep learning in the field of digital image processing, and the semantic segmentation model has been widely used. Semantic segmentation performs pixel-level labeling with a set of object categories (e.g., human, car, tree, and sky) for all image pixels; thus, it is generally a harder undertaking than image classification, which predicts a single label for the entire image [9]. Scholars have conducted in-depth research on U-Net and achieved good results. Ronneberger et al. [12] from the University of Freiburg proposed U-Net as a convolutional network that relies heavily on data augmentation to use the available annotated samples more efficiently and has a more significant advantage in the semantic segmentation of biomedical images. Kemker et al. [13] used convolutional neural networks for the semantic segmentation of 7-channel multispectral images to successfully recognize vegetation and calculate its coverage based on high-resolution aerial images of a park. Shojaiee and Baleghi [14] proposed a multispectral encoder fused aurous spatial pyramid pooling (EFASPP) U-Net to merge the features of visible and thermal images recorded in night traffic scenes. Yu et al. [15] used a U-Net model for semantic segmentation of RGB imagery in maize plants, and the results showed that the U-Net had an excellent segmentation accuracy for maize tassels under various complex situations. Wang et al. [16] took cucumber spectral images as the research object, and a U-Net was utilized to construct the semantic segmentation model for cucumber target leaf spot segmentation. Although extensive research has been carried out on U-Net, it is mostly used for a set of object classes, and no single study exists that adequately addresses color recognition. Furthermore, there is excellent potential for the semantic segmentation of painted images based on U-Net. The architecture consists of a contracting path to capture context and an expanding path for precise localization. U-Net relies on the strong use of data augmentation to learn more data features when the training samples are insufficient. However, when outputting semantic segmentation images for each color, it is easy to generate probabilistically unstable salt-and-pepper noise, which affects the segmentation accuracy. It is obvious that previous convolutional neural networks have limitations when semantic segmentation is performed on painted colors.

This paper presents a model called UR-Net to address the issue encountered by previous methods, which found it difficult to achieve color consistency when segmenting painted images. UR-Net adds an improved residual block to the contracting path and expanding path of U-Net for deeper training. UR-Net is additionally enhanced with the

ability to learn data features based on the original U-Net and perform accurate segmentation to achieve color consistency. This is achieved based on semantic segmentation and by making a finer division of the semantics so that each color is treated as a semantic category. Images for each color are obtained by training the UR-Net to recognize each color pixel in color painting samples. Compared with the previous convolutional neural networks, the optimized semantic segmentation model derived from the trained UR-Net in this paper is more accurate in recognizing color pixels. The proposed UR-Net is more applicable to painted images and can be used as a technical basis to achieve the color reconstruction of cultural color paintings.

2. Data Preparation

2.1. Color Painting Sample Acquisition and Processing

In this paper, three color paintings were selected as dataset samples for training and testing semantic segmentation. The painted images were collected using a Canon PowerShot SX100 IS digital camera (Canon Group, Tokyo, Japan) under a D65 illumination light source, shot in parallel with the color paintings, with a horizontal and vertical resolution of 180 dpi. The shooting parameters were as follows: aperture $f/4$, exposure time $1/160$ s, ISO 80, exposure compensation 0, focal length 6 mm, maximum aperture 2.96875, and flash turned off. We ensured the color accuracy of the images by controlling the camera type, spectrum of lighting, light intensity, etc. The images in the database were in jpg format. In Figure 1, the image of sample (a) is a drawing copy of the Dunhuang mural “Dancing Celestial Beings Playing Music” with a resolution of 938×1718 pixels. The image of sample (b) is “Meditating Monk” with a resolution of 500×744 pixels, from Cave 285 of the Dunhuang Mogao Caves, made during the Western Wei Dynasty. The image of sample (c) is “Celestial Beings” with a resolution of 468×725 pixels, from Cave 249 of the Dunhuang Mogao Caves, made during the Western Wei Dynasty.



Figure 1. Collected color painting samples. (a) Dancing Celestial Beings Playing Music. (b) Meditating Monk. (c) Celestial Beings.

However, the earth red pixels in sample (a) differ significantly from the actual earth red pixels. To verify the accuracy of the proposed model in recognizing earth red pixels, the color information of the actual earth red pigments is mapped to training sample (a) in this paper. First, pure earth red pigment is applied on cardboard. Then, the spectral reflectance factor and RGB information of the earth red pigment is obtained using an X-rite Ci64 integrating sphere spectrophotometer. Finally, the color information (spectral reflectance and RGB) of the earth red pigment is mapped to training sample (a) using the sample label as a mask. The mapping processing is as shown in Figure 2. By pigment mapping, the original color pixels of the color painting can be replaced with earth red pixels for better deep learning.

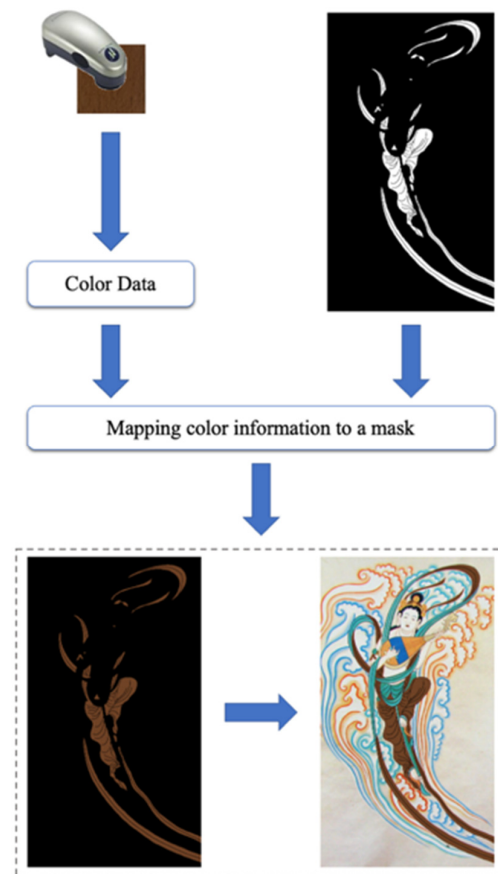


Figure 2. Mapping earth red pixels to sample a.

After the samples were collected, the original three channels of the R, G, and B datasets of the images were expanded to six channels of R, G, B, H, S, and V datasets with the D65 light source information. This can ensure that more information is learned from the samples during training. In addition, the RGB and H, S, and V channel images of the three samples are shown in Figures 3 and 4. The earth red and ultramarine blue pixels in test sample (c) are segmented by training on the earth red pixels in sample (a) and on the ultramarine blue pixels in sample (b).



Figure 3. RGB images of the color painting samples. (a) RGB image of a training sample. (b) RGB image of a training sample. (c) RGB image of a test sample.

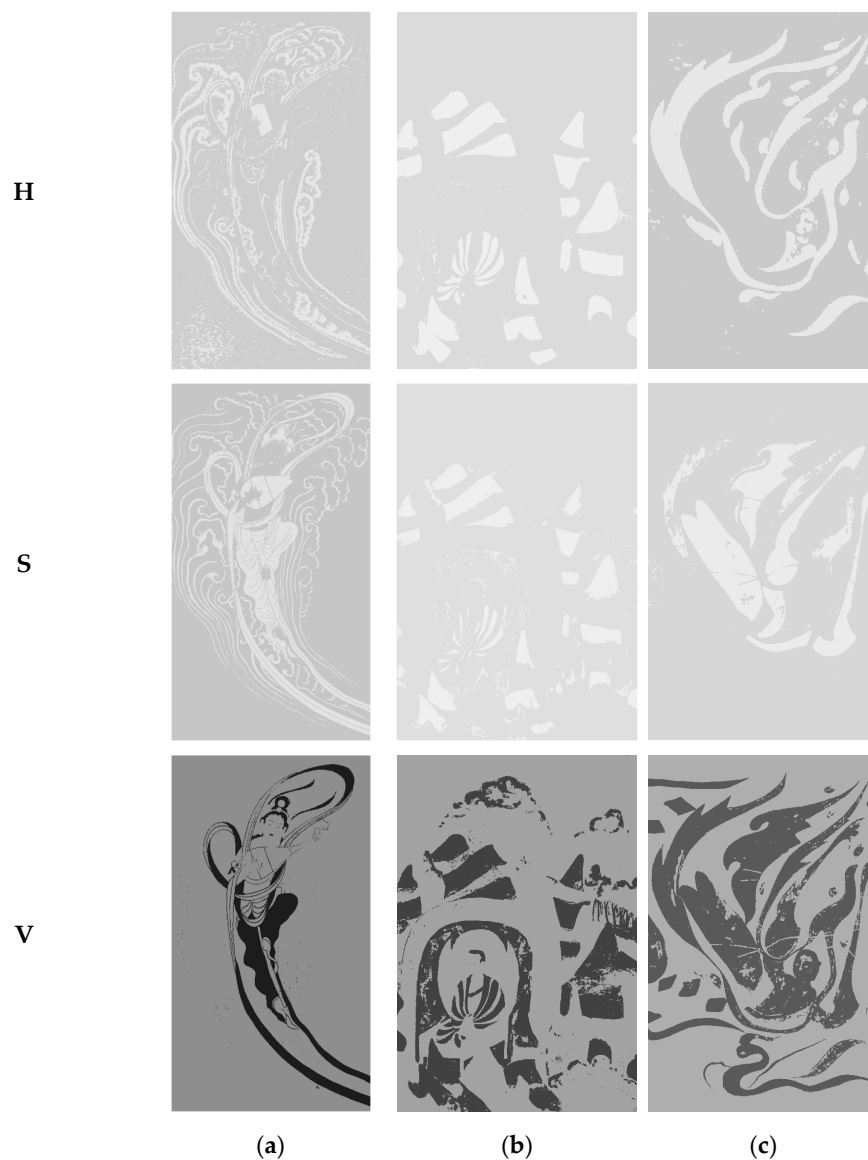


Figure 4. H, S, and V images of the color painting samples. (a) H, S, and V images of a training sample. (b) H, S, and V images of a training sample. (c) H, S, and V images of a test sample.

2.2. Color Painting Sample Labels

The color pixels in painted images need to be labeled before performing semantic segmentation. The desired color pixels are first segmented using the K-means clustering algorithm, and then the pixels are labeled using the Mask tool in MATLAB_R2022b. However, the K-means clustering algorithm is prone to overfitting and overidentifying color pixels for painted images, so it is necessary to combine the images and datasets of the three channels of H, S, and V in the samples to remove the overfitted pixels. It is also confirmed that a simple segmentation using color thresholding struggles to achieve color consistency when segmenting painted images with a high color gradient. The RGB images of the color paintings and the corresponding labeled results are shown in Figure 5. The blank areas in training sample (a) represent earth red, with 125,380 pixels, and the coverage of earth red is 7.64%. The blank areas in training sample (b) represent ultramarine blue, with 42,213 pixels, and the coverage of ultramarine blue is 11.35%. The blank areas in test sample (c) represent earth red (left), with 46,546 pixels, and the coverage of earth red is 13.72%. The blank areas in test sample (c) represent ultramarine blue (right), with 48,245 pixels, and the coverage of ultramarine blue is 14.22%. Labels were used for the semantic segmentation of supervised

learning, which can link the RGB values of each component pigment. This works for whatever two or more pigments were mixed.

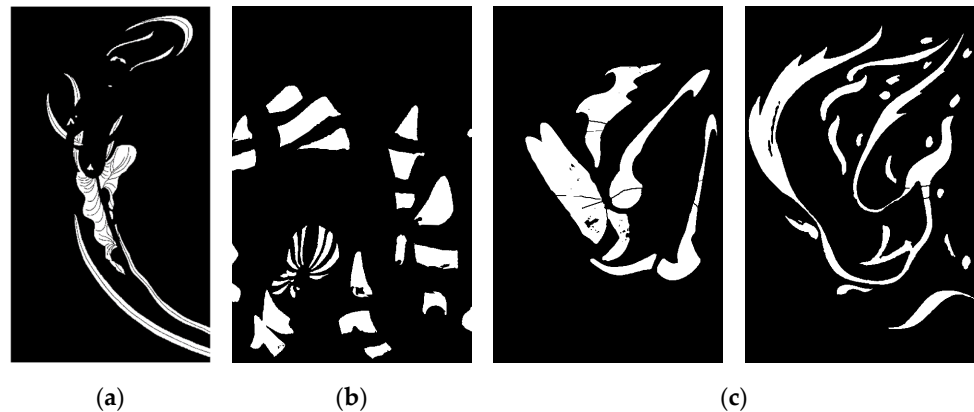


Figure 5. Labels of samples. (a) Earth red label of the training sample. (b) Ultramarine blue label of the training sample. (c) Earth red (left) label and ultramarine blue (right) label of the test sample.

3. Methodology

3.1. K-Means Clustering Algorithm

First, the algorithm selects K points from a sample image as the initial clustering centers, estimated as $K = 7$ in this paper. Second, the distance from each sample to the cluster is calculated, and the samples are classified into the nearest cluster center. Then, new cluster centers are calculated based on the classified samples. Finally, the above steps are repeated until the clustering centers of the two adjacent iterations no longer change. When using the K-means clustering algorithm to segment painted images, it is necessary to transform the images to the Lab color space to enhance the robustness of the algorithm [17]. However, the K-means clustering algorithm still has the two following limitations: (1) The number of initial clustering centers, K , needs to be given to determine the initial division. However, K is usually difficult to estimate. (2) Noise and isolated points significantly impact the segmentation results. This makes it difficult for K-means to achieve color consistency and reach the desired accuracy when performing full-cluster segmentation with complex painted images as input.

3.2. Residual Network

In convolutional neural networks, the accuracy of the training dataset rises and gradually saturates as the network depth increases. However, when the network is deepened again, the accuracy of the training dataset degrades rapidly. Thus, network degradation occurs, which makes it difficult for the samples to take advantage of the powerful feature extraction ability of the deep network during training. Such degradation is not caused by overfitting, and adding more layers to a suitably deep model leads to higher training error. ResNet is built from a series of residual blocks, and its main aim is to address network degradation [18]. A basic residual block can divide the workflow of ResNet into identity mapping and residual mapping steps, as shown in Figure 6.

The residual block consists of the weight layer, batch normalization (BN) layer, ReLU activation layer, and an additional layer as shown in Figure 6. The green box indicates the identity mapping, and the red box indicates the residual mapping. A residual block will output X_l as X_{l+1} after two mappings. The specific output of a base residual block is represented by Equation (1).

$$X_{l+1} = X_l + F(X_l + W_l) \quad (1)$$

where X_l is the input feature of the l -th residual block unit, W_l is the features learned of the l -th features learned, X_{l+1} is the output feature of the l -th residual block unit, and $F(X_l + W_l)$ is a residual mapping function.

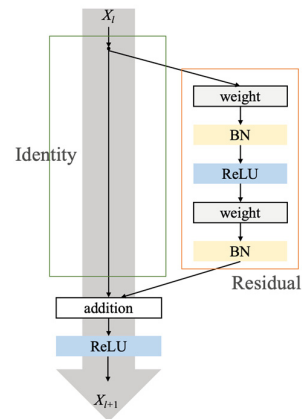


Figure 6. A residual block.

ResNet first yields the input X_l by its unique residual blocks and implements a function transformation on this input using residual mapping, which is described as $F(X_l + W_l)$. Then, the original input is added directly to this transformation using identity mapping, which is described as $x_l + F(X_l + W_l)$. Lastly, it outputs x_{l+1} . Because of its unique residual block structure, the network can obtain the feature information of the previous layer, which enhances the feature extraction ability of the network. However, similar to previous convolutional neural networks, ResNet needs to use a large number of samples for feature training to obtain a more accurate semantic segmentation model. As the creation of ancient painters, a color painting is usually unique. Hence, the number of samples is sparse, which easily leads to poor training results and makes it difficult to achieve color consistency when performing semantic segmentation.

3.3. U-Net Convolutional Network

To address the problem of sparse color painting samples, this paper constructs a network model based on the U-Net architecture and applies it to the semantic segmentation of painted colors. U-Net relies on a unique contracting path to capture the context of images and relies on a unique expanding path to achieve precise localization. Ronneberger, Fischer, and Brox [12] presented a network with a training strategy that relies on strong data augmentation to use the available annotated samples more efficiently and uses elastic deformations for data augmentation. This can learn more data features even if the number of training samples is sparse. U-Net inherits the advantages of the convolutional neural network while addressing its limitations and solving the problem of sparse samples. In the segmentation of painted images, it is possible to achieve color consistency and perform more accurate segmentation. The overall architecture has a symmetric ‘U’ shape, as shown in Figure 7. The completed U-Net with 58 layers and 61 connections is constructed.

The U-Net samples the input image and outputs a semantic segmented image by convolution and an activation function (ReLU). The feature extraction part is a contracting path for capturing the context of the image, in which each max-pooling layer decreases the image size. The upsampling part is an expanding path for precise localization, in which each upconvolution crops and fuses the extracted features to recover the image size. Finally, the skip connection allows the retrieved image to fuse more features so that the segmented image retains more detailed information.

When a U-Net is applied to paint image segmentation, although it can achieve color consistency and high segmentation accuracy, the output semantic segmentation image contains salt-and-pepper noise, and the probability of salt-and-pepper noise appears to be different for each color, so it cannot be applied to practical work. Therefore, it is necessary to improve the U-Net architecture to ensure segmentation accuracy while ensuring the output semantic maps of color paintings are free of noise.

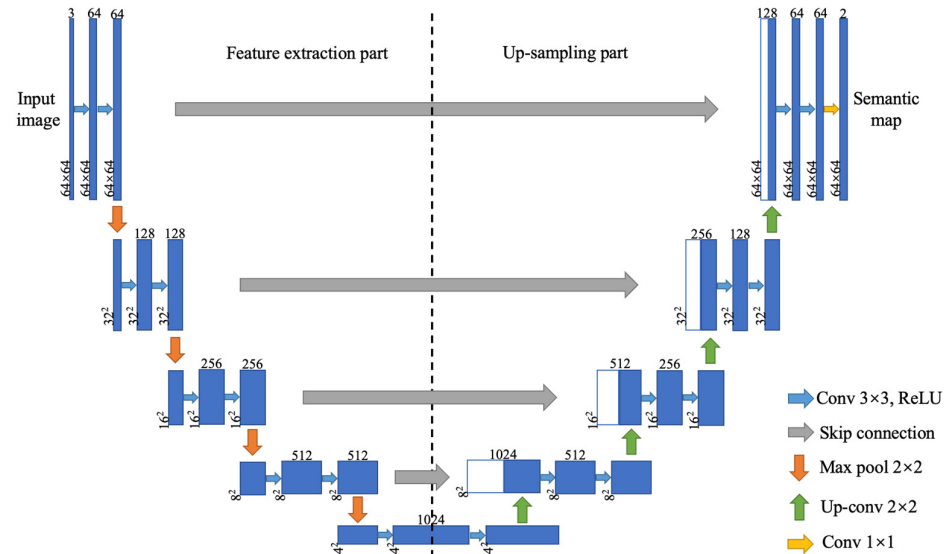


Figure 7. Schematic diagram of the U-Net architecture.

3.4. An Optimized UR-Net Model Based on the Combination of U-Net and Residual Blocks

ResNet can reduce the loss of feature information and perform deeper learning by using unique residual blocks when executing network training. U-Net can learn more feature information and perform more accurate semantic segmentation when the training sample data are sparse by relying on the strong use of data augmentation. However, ResNet requires a large number of training samples, and U-Net is prone to salt-and-pepper noise when performing semantic segmentation on painted images, ultimately leading to unsatisfactory segmentation results. Compared with previous networks, U-Net can learn more feature information during training but still faces the problem of information loss, which can easily lead to the output image containing salt-and-pepper noise. Therefore, the U-Net architecture is optimized in this paper by adding residual blocks to reduce the loss of feature information during training and increase the feature extraction ability of U-Net again on its original basis to segment painted images more accurately. First, the residual blocks are improved to fit the U-Net architecture. Second, residual blocks are added to the contracting and expanding paths of U-Net. Finally, the optimized UR-Net model is constructed for the semantic segmentation of color paintings.

Since the base residual block needs to be back-connected to the ReLU activation layer, adding it directly will result in a redundant ReLU activation layer in the U-Net structure, changing the original structure of the U-Net. As shown in Figure 8, the earth red pixels segmented by integrating the underlying residual block directly into the U-Net structure are still poorly segmented and have more pronounced noise.



Figure 8. The earth red pixels segmented by integrating the underlying residual block directly into the U-Net structure.

The base residual block is not well adapted to the U-Net architecture, so it is necessary to improve the residual block based on its original structure by moving the ReLU activation layer from under the additional layer to the residual mapping, as shown in Figure 9. This operation allows for the more efficient use of the residual block without changing the basic architecture of U-Net or the function of the residual block so that the residual block can be better integrated into the U-Net architecture.

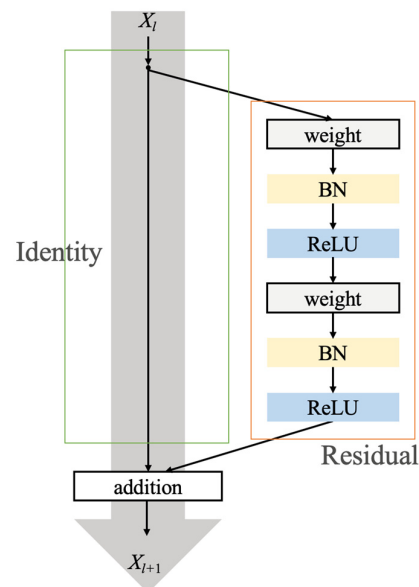


Figure 9. An improved residual block.

Then, the UR-Net needs to be constructed by adding the improved residual blocks in Figure 9 to the U-Net architecture in Figure 7, as shown in Figure 10. The completed UR-Net with 128 layers and 141 connections is constructed. This architecture applies the U-Net feature extraction and upsampling functions while reducing information loss using residual blocks. The UR-Net aims to remove the salt-and-pepper noise generated after U-Net segmentation, improve segmentation accuracy, and achieve color consistency.

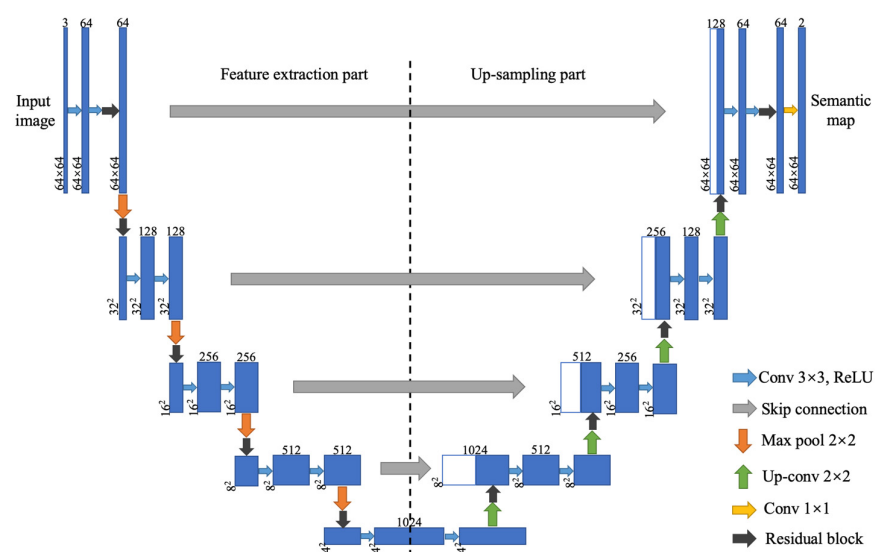


Figure 10. Schematic diagram of the UR-Net architecture.

In neural networks, the loss function is used to measure the degree of inconsistency between the predicted and true values of the model. The smaller the value of the loss

function is, the better the model is trained and the higher its robustness. The original U-Net architecture was applied to cell segmentation using a weighted cross-entropy loss function to measure the model parameters with the aim of strengthening the segmentation line between similar adherent cells by adding the corresponding weight proportion to the cross-entropy loss function. However, the boundaries between different colors are more obvious, and weighting the cross-entropy loss function might instead affect the learning rate and risk overfitting. Therefore, this paper directly uses the cross-entropy loss function to ensure that the training results of U-Net continue to approximate the true value, which enhances the accuracy and robustness of segmentation while accelerating the learning rate of the network.

The UR-Net output is followed by a softmax function, i.e., the predicted value is taken as the softmax output for cross-entropy calculation. The softmax function is represented by Equation (2).

$$f(x_i) = \frac{\exp(x_i)}{\sum_k \exp(x_k)} \quad (2)$$

where x_i is category i of sample x , x_k is all k categories in sample x , and $f(x_i)$ is the predicted probability.

The difference between the predicted probability and true probability determines the loss. The cross-entropy loss function calculated based on softmax is represented by Equation (3).

$$L = -\sum_{i=1}^n y_i \log f(x_i) \quad (3)$$

where y_i is the true probability, $f(x_i)$ is the predicted probability calculated by the softmax function, and L is the loss result of the variability between the true and predicted values.

The gradient $\partial_{x_i} L$ of L is the difference between the predicted probability and true probability, represented by Equation (4).

$$\partial_{x_i} L = \text{softmax}(x)_i - y_i \quad (4)$$

The specific workflow of the proposed UR-Net is shown in Figure 11. First, the training patches (training samples and labels) are input into the UR-Net architecture for training to obtain the trained UR-Net semantic segmentation model. Second, the datasets of painted images are input into the trained UR-Net for segmentation. Finally, semantically segmented images with different colors are obtained and output. These steps are repeated using ResNet and U-Net, and the segmentation results are derived and compared with those of UR-Net. The network training strategy and training parameter set are described in the following two subsections.

3.4.1. Network Training Strategy

Network training is the most important and tedious step in the whole segmentation process. We use a random patch extraction datastore to feed the training data to the network and thus perform data augmentation. This differs from the use of elastic deformations for data augmentation by Ronneberger, Fischer, and Brox [12]. This datastore extracts multiple corresponding random patches from an image datastore and pixel label datastore that contain ground-truth images and pixel label datasets. Patching is a crucial technique that effectively increases the available training data, and UR-Net relies strongly on data augmentation to use the available annotated patches more efficiently. The training patches (training samples and labels) are trained using the constructed UR-Net with the aim of obtaining the trained UR-Net semantic segmentation model. The image size of training sample (a) is 938×1718 pixels with six channels (R, B, G, H, S, V), including RGB and HSV color space. The image size of training sample (b) is 555×815 pixels with six channels (R, B, G, H, S, V), including RGB and HSV color space. In this paper, the patch input size of the network is 64×64 pixels regardless of the original patch size. After the input layer of UR-Net receives the training patches, it is scanned sequentially from left to right

and from top to bottom according to the 64×64 pixel window, and a high-resolution patch is divided into several low-resolution patches for training to obtain more feature information. This strategy has the two following advantages: first, when the patch size is greater than 64×64 pixels, it is not necessary to change the size in the input and output layers, which in turn simplifies the steps; second, data augmentation is performed by generating duplicates during patch segmentation. As shown in Figure 12, scanning a 64×125 pixel image in a 64×64 pixel window will result in the repeated recognition of 64×3 pixels of data information, allowing for data augmentation to be performed again to obtain more information for learning.

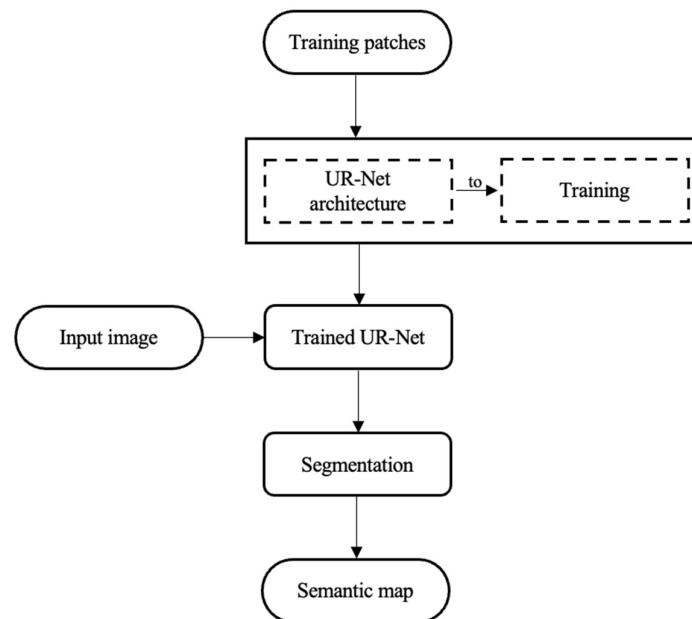


Figure 11. Workflow of the proposed UR-Net model.

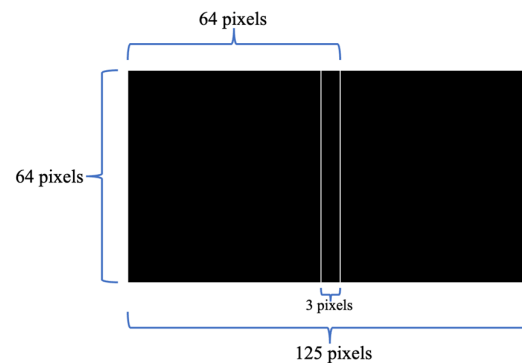


Figure 12. Repetition that may occur during the size segmentation process for patches.

3.4.2. Training Parameter Settings

In the training settings for samples, the maximum number of epochs is set to 15, and the mini-batch size is set to 4. This means that each mini-batch contains 8 patches of size 64×64 pixels. In each iteration, 30 mini-batches are extracted, and the training samples are divided into 120 patches for a total of 450 iterations. The training network is optimized using stochastic gradient descent with momentum (SGDM). The initial learning rate is set to 0.05, the momentum is set to 0.9, the L2 regularization is set to 0.0001, the gradient threshold is set to 0.05, and the verbose frequency is set to 20. Next, the same settings as above were used in training UR-Net, ResNet, and U-Net. One set of training progress data for UR-Net on training sample (b) is shown in Table 1.

Table 1. Training progress.

Epoch	Iteration	Time Elapsed (m:s)	Mini-Batch Accuracy	Mini-Batch Loss
1	1	0:01	05.49%	10.288
1	20	0:08	98.46%	0.0407
2	40	0:15	98.66%	0.0302
2	60	0:21	99.11%	0.0224
3	80	0:27	98.59%	0.0322
4	100	0:33	99.52%	0.0130
4	120	0:40	99.32%	0.0164
5	140	0:46	98.68%	0.0332
6	160	0:53	99.43%	0.0138
6	180	0:59	99.62%	0.0093
7	200	1:05	97.65%	0.0740
8	220	1:12	99.50%	0.0118
8	240	1:18	99.22%	0.0196
9	260	1:24	99.63%	0.0093
10	280	1:31	99.30%	0.0167
10	300	1:37	99.76%	0.0062
11	320	1:44	99.78%	0.0054
12	340	1:50	99.77%	0.0053
12	360	1:57	99.64%	0.0090
13	380	2:03	99.81%	0.0049
14	400	2:09	99.61%	0.0069
14	420	2:16	99.60%	0.0085
15	440	2:22	99.32%	0.0158
15	450	2:25	99.76%	0.0077

All experiments were implemented in MATLAB R2020a with an Intel i7-770HQ 2.81 GHz CPU, 8 GB RAM, NVIDIA GeForce GTX 1050Ti graphics card, and 4 GB VRAM. The results show that UR-Net can complete the training after 450 iterations in 145 s with a mini-batch accuracy of 99.81% and a mini-batch loss of 0.0049; thus, the network inferences converge to a color mapping.

4. Experimental Evaluation

To explore the performance of the proposed semantic segmentation model, which is based on UR-Net for the color segmentation of painted images, a series of experiments were conducted using the collected color painting dataset by comparing the segmentation results of the ResNet model and the U-Net model.

4.1. Evaluation Metrics

Based on the confusion matrix [15], the semantic segmentation accuracy is quantified by calculating evaluation metrics, such as pixel accuracy, recall, precision, F1-score, and intersection over union [19,20]. The confusion matrix is the classification result of the statistical classification model. It indicates the numbers of samples correctly and incorrectly categorized, as shown in Table 2. The blank areas denote the correctly categorized pixels, and other areas denote the incorrectly categorized pixels in the label. Whether the segmentation is manual or computerized (e.g., clustering algorithms or semantic segmentation), the segmented image areas cannot be guaranteed to be 100% correct. In this paper, a correct segmentation criterion is given by the labels, and then the correct areas in the labels are evaluated against the segmented areas and their differences are compared.

Table 2. Confusion matrix.

Confusion Matrix		Actual Class	
		Label	Background
Predicted class	label	True Positives (TP)	False Positives (FP)
	Background	False Negatives (FN)	True Negatives (TN)

In the dataset, the ratio of the number of correctly segmented pixels (TP + TN) to the total number of pixels (TP + FN + FP + TN) is the pixel accuracy (PA), which is used to globally evaluate the percentage of correctly segmented pixels, represented by Equation (5).

$$PA = \frac{TP + TN}{TP + FN + FP + TN} \quad (5)$$

The F1-score is the summed average of recall and precision. Recall is the proportion of correctly segmented label pixels (TP) among all actual label samples (TP + FN), and precision is the proportion of correctly segmented label pixels (TP) among all predicted label samples (TP + FP), represented by Equations (6) and (7).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

After applying Equations (6) and (7), the F1-score can be calculated, which indicates the matching quality of the predicted boundary of each class with its actual boundary on all images and evaluates the boundary accuracy, represented by Equation (8).

$$F1 - \text{score} = \frac{2 \cdot \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

The intersection over union (IoU) is a standard metric for evaluating the accuracy of semantic segmentation and indicates the proportion of the overlap area between the predicted and actual segmentation to the joint area, indicating the regional smoothness, represented by Equation (9).

$$IoU = \frac{TP}{TP + FP + FN} \quad (9)$$

In Equations (5)–(9) and Table 2, TP is the number of correctly segmented label pixels, which are true positives. TN is the number of correctly segmented background pixels, which are true negatives. FN is the number of unrecognized label pixels, which are false negatives. FP is the number of unrecognized background pixels, which are false positives.

4.2. Ablation Experiments

Residual blocks in different quantities and with different structures, added to different positions of the U-Net framework, will produce different output effects. In order to explore the effect of the number of residual blocks and the advantage of the improved residual blocks, the ablation experiment shown in Figure 13 is set up to add two different structures of residual blocks (basic residual blocks and improved residual blocks) to the I, II, III, IV, and V positions of the U-Net framework, and to explore the UR-Net network structure that is the most suitable for the color images as the semantic segmentation model. Since the constructed U-Net framework is left-right symmetric, the added positions are also left-right symmetric.

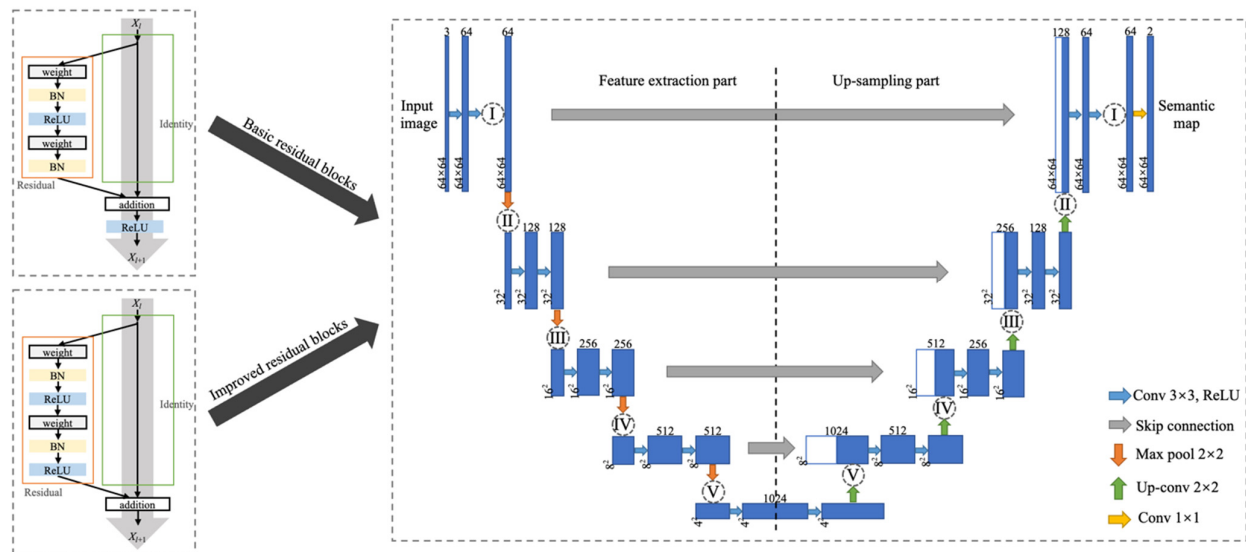


Figure 13. UR-Net structure setup.

The adding positions of the residual blocks are set as follows:

Add two basic residual blocks to position I of the U-Net framework, denoted as UR-Net-2b. Add two improved residual blocks to position I of the U-Net framework, denoted as UR-Net-2i.

Add four basic residual blocks to positions I and II of the U-Net framework, denoted as UR-Net-4b. Add four improved residual blocks to positions I and II of the U-Net framework, denoted as UR-Net-4i.

Add six basic residual blocks to positions I, II, and III of the U-Net framework, denoted as UR-Net-6b. Add six improved residual blocks to positions I, II, and III of the U-Net framework, denoted as UR-Net-6i.

Add eight basic residual blocks to positions I, II, III, and IV of the U-Net framework, denoted as UR-Net-8b. Add eight improved residual blocks to positions I, II, III, and IV of the U-Net framework, denoted as UR-Net-8i.

Add 10 basic residual blocks to positions I, II, III, IV, and V of the U-Net framework, denoted as UR-Net-10b. Add 10 improved residual blocks to positions I, II, III, IV, and V of the U-Net framework, denoted as UR-Net-10i.

The two types of residual blocks are added in different numbers to different positions in the network structure, as shown in Table 3; “✓” indicates the selected positions and numbers in the network structure.

Table 3. Positions of basic and improved residual blocks.

Basic	Improved	Number	Positions				
			I	II	III	IV	V
UR-Net-2b	UR-Net-2i	2	✓				
UR-Net-4b	UR-Net-4i	4	✓	✓			
UR-Net-6b	UR-Net-6i	6	✓	✓	✓		
UR-Net-8b	UR-Net-8i	8	✓	✓	✓	✓	
UR-Net-10b	UR-Net-10i	10	✓	✓	✓	✓	✓

Each UR-Net with a different number of residual blocks was trained according to the set training methods and parameters, and the network structure with the highest accuracy was selected to be used as UR-Net. IoU and other data were calculated using Equations (5)–(9) to validate the usefulness of the 10 UR-Net structures, and the results are shown in Table 4. The calculation results show that the IoU and other data of UR-Net-10i are

higher than the other model structures, i.e., the maximum accuracy can be achieved when 10 improved residual blocks are added to the U-Net structure. The semantic segmentation of earth-red pixels using the model shown in Figure 14 still produces salt-and-pepper noise, which further verifies the advantage of the improved residual blocks. Therefore, UR-Net-10i is used as the UR-Net network model for semantic segmentation of colored images.

Table 4. Accuracy comparison of ten structures of UR-Net.

Structure	Dataset	PA	Recall	Precision	F1-Score	IoU
UR-Net-2b	Earth red	0.9891	0.9875	0.9364	0.9613	0.9254
	Ultramarine blue	0.9872	0.9845	0.9299	0.9564	0.9165
UR-Net-2i	Earth red	0.9890	0.9912	0.9329	0.9612	0.9253
	Ultramarine blue	0.9876	0.9764	0.9390	0.9573	0.9182
UR-Net-4b	Earth red	0.9894	0.9812	0.9437	0.9621	0.9269
	Ultramarine blue	0.9867	0.9578	0.9497	0.9537	0.9116
UR-Net-4i	Earth red	0.9883	0.9873	0.9314	0.9585	0.9203
	Ultramarine blue	0.9877	0.9840	0.9332	0.9580	0.9193
UR-Net-6b	Earth red	0.9882	0.9881	0.9301	0.9582	0.9198
	Ultramarine blue	0.9874	0.9600	0.9523	0.9561	0.9159
UR-Net-6i	Earth red	0.9888	0.9886	0.9336	0.9603	0.9237
	Ultramarine blue	0.9870	0.9851	0.9280	0.9557	0.9151
UR-Net-8b	Earth red	0.9889	0.9851	0.9371	0.9605	0.9240
	Ultramarine blue	0.9868	0.9738	0.9359	0.9545	0.9130
UR-Net-8i	Earth red	0.9882	0.9818	0.9356	0.9581	0.9196
	Ultramarine blue	0.9870	0.9532	0.9552	0.9542	0.9125
UR-Net-10b	Earth red	0.9875	0.9900	0.9244	0.9561	0.9159
	Ultramarine blue	0.9878	0.9769	0.9401	0.9581	0.9197
UR-Net-10i	Earth red	0.9906	0.9819	0.9510	0.9662	0.9346
	Ultramarine blue	0.9890	0.9670	0.9561	0.9615	0.9259

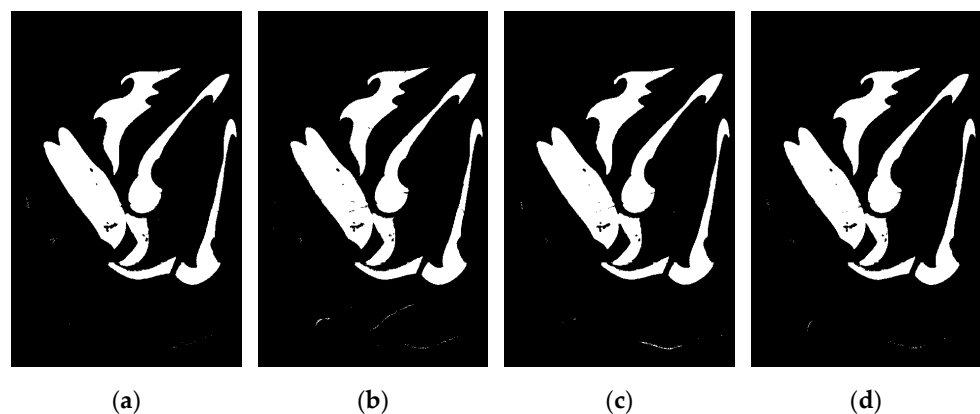


Figure 14. Salt-and-Pepper noise from different structural UR-Net segmentations. (a) UR-Net-2b. (b) UR-Net-10b. (c) UR-Net-2i. (d) UR-Net-6i.

4.3. Comparison Experiment

First, K-means, ResNet, U-Net, and the proposed UR-Net were used to semantically segment the earth red and ultramarine blue pixels in the color painting test samples, and the corresponding color segmentation images were obtained and compared with the true results, as shown in Figure 15. The number of K-means clusters is difficult to estimate, and the segmentation effect achieved for the color painting images is not ideal. When the sample dataset is sparse, the segmentation accuracy achieved by the semantic segmentation model trained by ResNet is low, and it cannot accurately recognize the color pixels, easily

causing overfitting. Although the semantic segmentation model trained by U-Net can achieve better color consistency, the output color semantic map still contains salt-and-pepper noise, which affects the segmentation accuracy. In this paper, the proposed UR-Net combines the advantages of ResNet and U-Net by adding residual blocks to the U-Net architecture, which make the network recognize pixels more accurately in the training process and remove noise while ensuring segmentation accuracy, thus achieving color consistency. From Figure 15, it can be seen that the segmentation accuracy of the UR-Net model is significantly higher than that of the K-means clustering algorithm and the ResNet model, and the generated salt-and-pepper noise is also considerably reduced compared with that of the U-Net model, so the proposed UR-Net model is more suitable for acquiring the semantic segmentation images of painted colors.

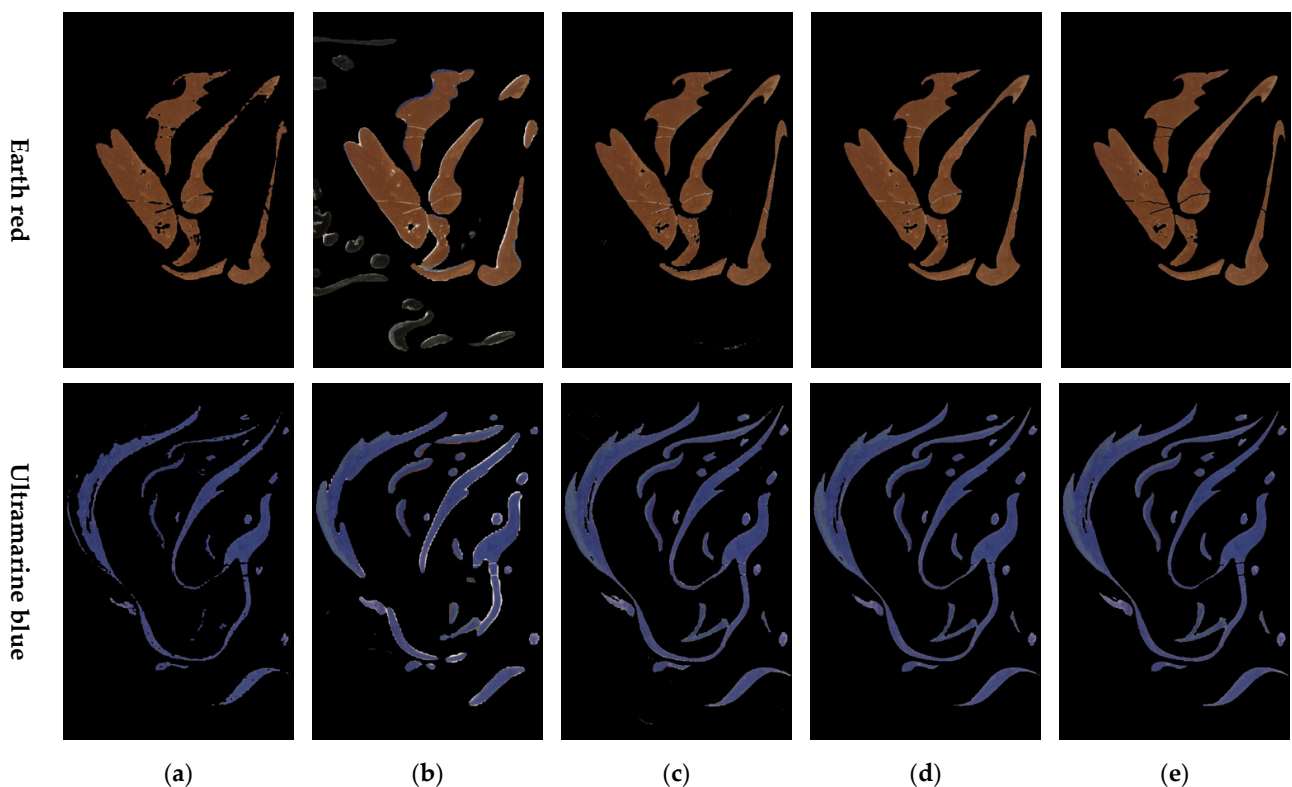


Figure 15. Comparison of different segmentation results with the true results. (a) K-means clustering algorithm. (b) ResNet. (c) U-Net. (d) The proposed UR-Net. (e) True results.

Second, the numbers of earth red pixels and ultramarine blue pixels in the segmentation results of the four methods are compared with the true results, as shown in Figure 16. By K-means segmentation, the number of earth red pixels is 42,794, and the number of ultramarine blue pixels is 32,881. By ResNet segmentation, the number of earth red pixels is 70,324, and the number of ultramarine blue pixels is 49,596. By U-Net segmentation, the number of earth red pixels is 48,277, and the number of ultramarine blue pixels is 51,118. By the proposed UR-Net segmentation, the number of earth red pixels is 48,058, and the number of ultramarine blue pixels is 48,981. The true number of earth red pixels is 46,546, and that of ultramarine blue pixels is 48,433. It is obvious that the segmentation results of the proposed method are more similar to the true results. However, the number of pixels recognized by the network model cannot be fully used as an evaluation criterion because the trained model may overfit during segmentation, which may cause it to overidentify color pixels and background pixels, resulting in inaccurate pixel numbers. In other words, if the model segmentation accuracy is higher, the number of recognized pixels must be more accurate. In contrast, if the model accuracy is low, the number of recognized pixels

may not differ much from the true value. Therefore, there is a need to use a more objective method to evaluate segmentation accuracy.

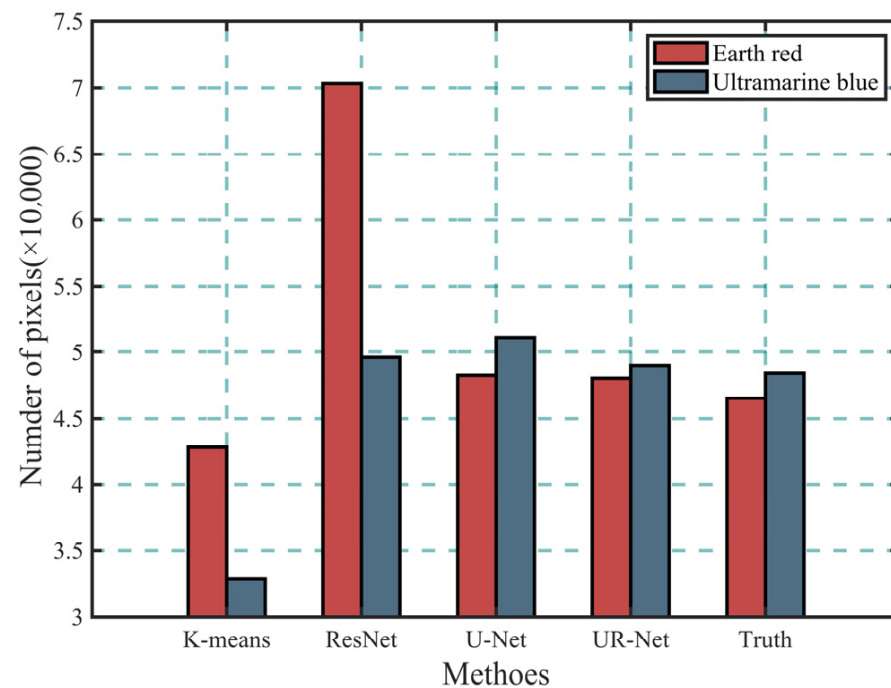


Figure 16. Numbers of pixels in different segmentation results compared to the true results.

Finally, the confusion matrix is calculated to quantify the difference between the true results and the segmentation results for a more objective evaluation. As shown in Figure 17, the blank area represents TP, the black area represents TN, the magenta area represents FN, and the green area represents FP. The magenta area indicates that the label pixels are recognized as background pixels, and the green area indicates that the background pixels are recognized as label pixels, both of which show the difference between the true results and the segmentation results.

The value of $TP + TN$ denotes the total number of correctly recognized pixels, and the value of $FN + FP$ denotes the total number of unrecognized pixels. If the value of $TP + TN$ is larger and the value of $FN + FP$ is smaller, the magenta and green areas are smaller. Comparing the four segmentation methods in Figure 17, the proposed UR-Net model resulted in the smallest magenta and green areas. The $TP + TN$ and $FN + FP$ values of the earth red and ultramarine blue pixels can be calculated from the TP, TN, FN, and FP values of the confusion matrix after the test sample segmentation is completed, as shown in Table 5. The results showed that the proposed UR-Net model has a $TP + TN$ value of 336,102 when segmenting the earth red pixels, which is larger than those of the other methods, and an $FN + FP$ value of 3198, which is smaller than those of the other methods. The $TP + TN$ value of 335,393, when segmenting the ultramarine blue pixels, is larger than those of the other methods, and the $FN + FP$ value of 3907 is smaller than those of the other methods. It is demonstrated that the proposed UR-Net model outperforms the other methods regarding the number of pixels correctly recognized in segmentation.

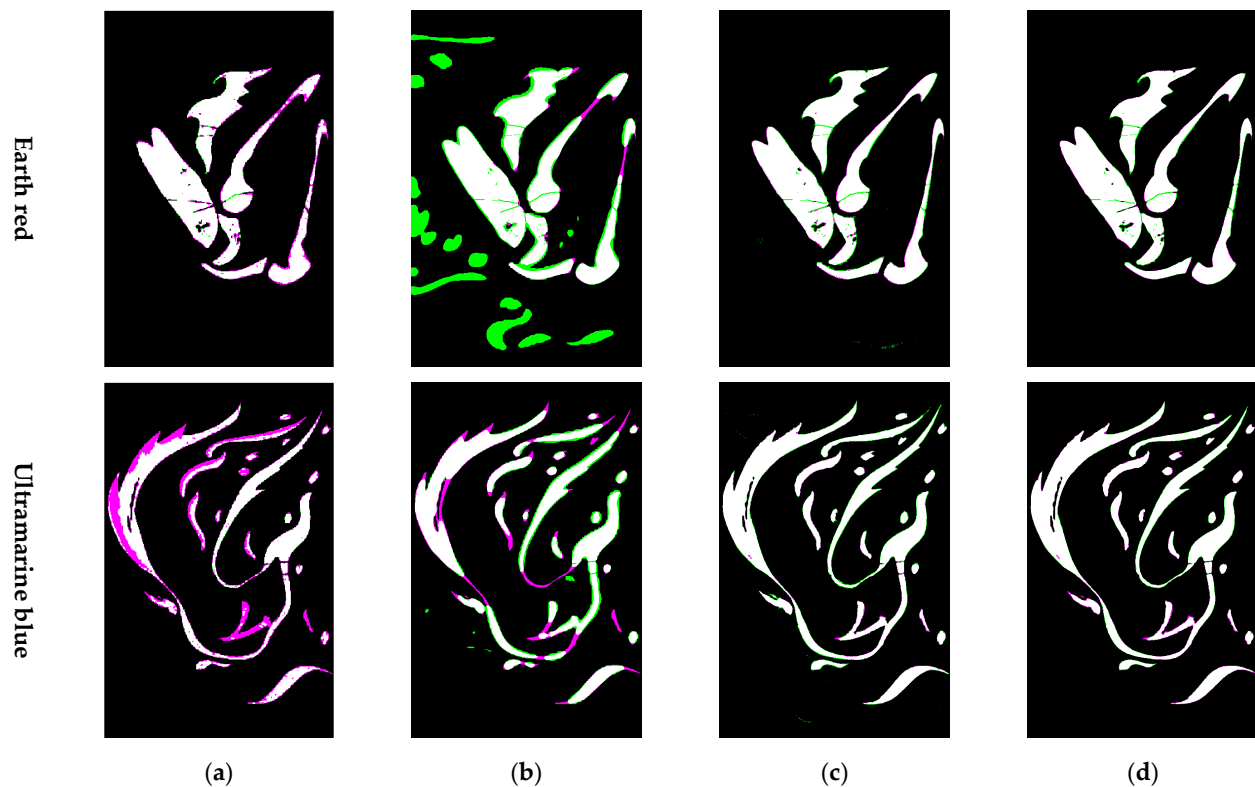


Figure 17. Differences between the true results and segmentation results for different segmentation methods. (a) K-means clustering. (b) ResNet. (c) U-Net. (d) The proposed UR-Net.

Table 5. TP + TN and FN + FP calculated based on TP, TN, FN, and FP.

Method	Dataset	TP	TN	FN	FP	TP + TN	FN + FP
K-means	Earth red	42,135	292,095	4411	659	334,230	5070
	Ultramarine blue	32,638	290,624	15,795	243	323,262	16,038
ResNet	Earth red	45,122	267,552	1424	25,202	312,674	26,626
	Ultramarine blue	41,922	283,193	6511	7674	325,115	14,185
U-Net	Earth red	45,180	289,657	1366	3097	334,837	4463
	Ultramarine blue	47,480	287,229	953	3638	334,709	4591
UR-Net	Earth red	45,703	290,399	843	2355	336,102	3198
	Ultramarine blue	46,833	288,719	1600	2148	335,552	3748

After determining the FN, FP, TP, and TN values in the confusion matrix, data such as PA, F1-score, and IoU can be calculated to evaluate the accuracy of segmentation, as shown in Table 4. The results are as follows: (1) The PA is 0.9906, the F1-score is 0.9662, and the IoU is 0.9346 when using the proposed UR-Net model to segment the earth red pixels. (2) The PA is 0.9890, the F1-score is 0.9615, and the IoU is 0.9259 when using the proposed UR-Net model to segment the ultramarine blue pixels. All the values are higher than those of the other segmentation methods, and the IoU values are all above 0.92, which gives a very high segmentation accuracy. Although the color paintings segmented by the U-Net model can also achieve high accuracy, there is salt-and-pepper noise, so this method cannot meet the requirements and cannot be applied to the segmentation of color paintings, as shown in Figure 17c. We have found salt-and-pepper noise in the images that were output by U-Net. Combining the data in Figure 3 and Table 6, it can be seen that although all three painted images belong to Dunhuang murals, there are significant differences in shape and color between the training samples and the test samples. In this context, the proposed UR-Net model can still recognize the earth red and ultramarine blue pixels more accurately, which

proves the high application value of UR-Net in the color recognition and segmentation of color paintings.

Table 6. Results of the accuracy evaluation.

Method	Dataset	PA	Recall	Precision	F1-Score	IoU
K-means	Earth red	0.9851	0.9052	0.9846	0.9433	0.8926
	Ultramarine blue	0.9527	0.6739	0.9926	0.8028	0.6705
ResNet	Earth red	0.9215	0.9694	0.6416	0.7722	0.6289
	Ultramarine blue	0.9582	0.8656	0.8453	0.8553	0.7472
U-Net	Earth red	0.9868	0.9707	0.9358	0.9529	0.9101
	Ultramarine blue	0.9865	0.9803	0.9358	0.9538	0.9118
UR-Net	Earth red	0.9906	0.9819	0.9510	0.9662	0.9346
	Ultramarine blue	0.9890	0.9670	0.9561	0.9615	0.9259

Combining the experimental data in Tables 3 and 4 and Figure 17, it can be seen that the proposed UR-Net model has a higher accuracy than the other two methods for segmenting the painting dataset and is more suitable for complex painted images. The segmentation results of the ultramarine blue pixels in Figure 17a show large magenta areas, indicating that the K-means clustering algorithm overfitted when performing full-cluster segmentation on the input image, overidentifying the background pixels and struggling to achieve color consistency. A good result is obtained when segmenting the earth red pixels, but the data values, including IoU, are still inferior to those of the proposed method, and the required accuracy is not achieved. The comparison images of earth red and ultramarine blue pixels in Figure 17b both show large green and magenta areas, and the data values, including IoU, are widely different from those expected, indicating that ResNet displays insufficient segmentation accuracy in achieving color consistency, which is mainly due to the insufficient amount of data learned by ResNet due to the small number of training samples. The green and magenta areas in Figure 17c are significantly reduced, and the data values such as IoU are also significantly improved, indicating that U-Net can still achieve better color consistency when performing segmentation tasks with insufficient training samples; however, the output image contains salt-and-pepper noise, which affects the segmentation accuracy. To address the above issues, the proposed UR-Net adds improved residual blocks to the U-Net architecture, relying on the powerful feature extraction ability of U-Net and the function of residual blocks to reduce information loss, which allows for more accurate training using a small number of samples. On the other hand, UR-Net displays some uncertainty during the training process and needs to be trained several times to obtain satisfactory results; therefore, we aim to improve this in future work.

4.4. Intuitive Evaluation

To evaluate the generalizability of the results, this paper uses UR-Net model to train the test samples again for the semantic segmentation of earth red and ultramarine blue pixels in painted images, as shown in Figure 18. The segmentation results show that the semantic segmentation model trained by UR-Net can achieve high-precision color segmentation and color consistency for both earth red and ultramarine blue pixels and for both faded color paintings and copied drawings.



Figure 18. Painted images segmented by UR-Net.

5. Conclusions

We have shown the usefulness of UR-Net for color recognition and segmentation when there is not enough annotated image data. Our network model has been shown to improve the semantic segmentation performance when compared to traditional networks and unsupervised feature extraction techniques. Our research adds the improved residual blocks of ResNet to the U-Net architecture to optimize the model and constructs a UR-Net for the semantic segmentation of color paintings. The experiment was conducted with earth red and ultramarine blue and can theoretically identify any pigments in the color painting. First, residual blocks were constructed and improved to make them suitable for the U-Net architecture. Second, the improved residual blocks were added to the U-Net architecture to construct UR-Net. Then, training samples were trained using UR-Net to obtain a semantic segmentation model applicable to painted colors. Finally, the proposed model was used to segment the painted images to derive semantic maps of the corresponding colors. The segmentation accuracy of the proposed UR-Net is examined by calculating the parameters of PA, recall, precision, F1-score, and IoU and comparing the three methods of K-means, ResNet, and U-Net. The results show that UR-Net has the advantages of a short training period, accurate pixel recognition, and high segmentation accuracy, in addition to achieving color consistency without the need for a large number of samples for training, which is very suitable for color-rich painted images. When segmenting color painting images with the proposed method, only the color pixels in a color painting need to be labeled as samples for training beforehand, and then the rest of the samples can be segmented, improving efficiency and ensuring segmentation accuracy.

Author Contributions: Conceptualization, S.F. and Z.L.; methodology, S.F.; software, S.F.; validation, S.F., L.L. and S.L.; formal analysis, Z.L.; investigation, S.F.; resources, S.F.; data curation, S.F.; writing—original draft preparation, S.F.; writing—review and editing, S.F.; visualization, S.F.; supervision, S.L.; project administration, Z.L.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: National Social Science Fund (Grant No. 21CTQ041); Natural Science Foundation of Shandong Province (Grant No. ZR2024MG044); and Ministry of Education of Humanities and Social Science Project (Grant No. 24YJAZH095).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: The author would like to thank their mentor for his help and the reviewers for their insightful comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Berns, R.S. Digital color reconstructions of cultural heritage using color-managed imaging and small-aperture spectrophotometry. *Color Res. Appl.* **2019**, *44*, 531–546. [\[CrossRef\]](#)
- Sun, F.; Wang, R.; Liang, Y.; Liu, J.; Yao, X.; Zhao, F. Analysis of Color Paint for Moyan Stone Carvings in Shuining Temple in Bazhong. *Spectrosc. Spectr. Anal.* **2022**, *42*, 505–511.
- Liang, J.; Wan, X. Non-Destructive Pigment Identification Method of Ancient Murals Based on Visible Spectrum. *Spectrosc. Spectr. Anal.* **2017**, *37*, 2519–2526.
- Kirchner, E.; van der Lans, I.; Ligterink, F.; Geldof, M.; Megens, L.; Meedendorp, T.; Pilz, K.; Hendriks, E. Digitally reconstructing Van Gogh's Field with Irises near Arles part 3: Determining the original colors. *Color Res. Appl.* **2018**, *43*, 311–327. [\[CrossRef\]](#)
- Berns, R.S.; Byrns, S.; Casadio, F.; Fiedler, I.; Gallagher, C.; Imai, F.H.; Newman, A.; Taplin, L.A. Rejuvenating the color palette of Georges Seurat's A Sunday on La Grande Jatte—1884: A simulation. *Color Res. Appl.* **2006**, *31*, 278–293. [\[CrossRef\]](#)
- Hartigan, J.A.; Wong, M.A. A k-means clustering algorithm. *Applied statistics. Algorithms* **1978**, *1326*, 100–108.
- Docquier, P.L.; Paul, L.; Menten, R.; Cartiaux, O.; Francq, B.; Banse, X. Measurement of bone cyst fluid volume using k-means clustering. *Magn. Reson. Imaging* **2009**, *27*, 1430–1439. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liang, J.; Liu, A.; Zhou, J.; Xin, L.; Zuo, Z.; Liu, Z.; Luo, H.; Chen, J.; Hu, X. Optimized method for segmentation of ancient mural images based on superpixel algorithm. *Front. Neurosci.* **2022**, *16*, 1031524–1031536. [\[CrossRef\]](#) [\[PubMed\]](#)
- Minaee, S.; Boykov, Y.Y.; Porikli, F.; Plaza, A.J.; Kehtarnavaz, N.; Terzopoulos, D. Image Segmentation Using Deep Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3523–3542. [\[CrossRef\]](#) [\[PubMed\]](#)
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- Yang, Y.; Wang, L.; Huang, M.; Zhu, Q.; Wang, R. Polarization imaging based bruise detection of nectarine by using ResNet-18 and ghost bottleneck. *Postharvest Biol. Technol.* **2022**, *189*, 111916–111927. [\[CrossRef\]](#)
- Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
- Kemker, R.; Salvaggio, C.; Kanan, C. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 60–77. [\[CrossRef\]](#)
- Shojaee, F.; Baleghi, Y. EFASPP U-Net for semantic segmentation of night traffic scenes using fusion of visible and thermal images. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105627–105641. [\[CrossRef\]](#)
- Yu, X.; Yin, D.; Nie, C.; Ming, B.; Xu, H.; Liu, Y.; Bai, Y.; Shao, M.; Cheng, M.; Liu, Y.; et al. Maize tassel area dynamic monitoring based on near-ground and UAV RGB images by U-Net model. *Comput. Electron. Agric.* **2022**, *203*, 107477–107491. [\[CrossRef\]](#)
- Wang, X.; Li, H.; Lv, L.; Han, D.; Wang, Z. Segmentation of Cucumber Target Leaf Spot Based on U-Net and Visible Spectral Images. *Spectrosc. Spectr. Anal.* **2021**, *41*, 1499–1504.
- Zhao, Y. Image Segmentation and Pigment Mapping of Cultural Heritage Based on Spectral Imaging. Ph.D. Thesis, Rochester Institute of Technology, New York, NY, USA, 2008.
- He, K.; Zhang, X.; Ren, S.; Sun, J. Identity Mappings in Deep Residual Networks. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 630–645.
- Wu, J.; Yang, G.; Yang, H.; Zhu, Y.; Li, Z.; Lei, L.; Zhao, C. Extracting apple tree crown information from remote imagery using deep learning. *Comput. Electron. Agric.* **2020**, *174*, 105504–105518. [\[CrossRef\]](#)
- Everingham, M.; Eslami, S.M.A.; Van Gool, L.; Williams, C.K.I.; Winn, J.; Zisserman, A. The Pascal Visual Object Classes Challenge: A Retrospective. *Int. J. Comput. Vis.* **2014**, *111*, 98–136. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.