Robust Nonparametric Distribution Transfer with Exposure Correction for Image Neural Style Transfer

Shuai Liu * , **Caixia Hong** , **Jing He and Zhiqiang Tian**

School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China; hongcx@stu.xjtu.edu.cn (C.H.); smile_hj@stu.xjtu.edu.cn (J.H.); zhiqiangtian@mail.xjtu.edu.cn (Z.T.)

* Correspondence: sh_liu@mail.xjtu.edu.cn; Tel.: +86-1569-197-3161

check for updates

Abstract: Image neural style transfer is a process of utilizing convolutional neural networks to render a content image based on a style image. The algorithm can compute a stylized image with original content from the given content image but a new style from the given style image. Style transfer has become a hot topic both in academic literature and industrial applications. The stylized results of current existing models are not ideal because of the color difference between two input images and the inconspicuous details of content image. To solve the problems, we propose two style transfer models based on robust nonparametric distribution transfer. The first model converts the color probability density function of the content image into that of the style image before style transfer. When the color dynamic range of the content image is smaller than that of style image, this model renders more reasonable spatial structure than the existing models. Then, an adaptive detail-enhanced exposure correction algorithm is proposed for underexposed images. Based this, the second model is proposed for the style transfer of underexposed content images. It can further improve the stylized results of underexposed images. Compared with popular methods, the proposed methods achieve the satisfactory qualitative and quantitative results.

Keywords: robust nonparametric distribution transfer; exposure correction; neural style transfer

1. Introduction

Style transfer seeks an artistic style reproduction on daily photos, enabling the photos retain the original content while presenting a prescribed artistic style [1–3]. The artistic style generally includes the color style and the texture style. By convention, the daily photo in style transfer is called content image, while the image who renders the artistic style is the style image. Style transfer makes artistic creation no longer distant for most people [4]. The process of neural style transfer is showed in Figure 1. The traditional artistic stylization algorithms are related to an area called non-photorealistic rendering (NPR), most of which transfer particular artistic styles [1,4]. In recent years, the development of deep learning [5–7] has made a breakthrough in style transfer [8–11]. The neural network enables the extraction of artistic styles, which makes it easier to convert daily photos into artistic photos [12]. For example, Gatys et al. [8] use the convolutional neural network (CNN) to reconstruct the content and style, and optimize the stylized image iteratively based on a loss function. From then on, CNN-based neural style transfer has become a hot topic [9–11,13,14].

For a long time, the research on neural style transfer mainly focused on two issues. It is hoped that the stylized images have a strong artistic sense and the model has good generalization for a wide application meanwhile. However, the application of style transfer is affected by many factors. Style images with large color differences from content images often generate errors during style transfer, while the inconspicuous details in low-illumination images also lead to unsatisfactory results.
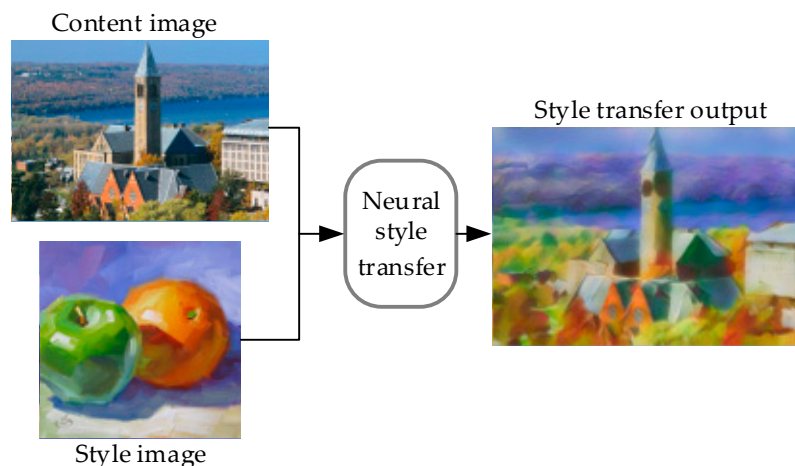
Figure 1. The processing of neural style transfer.

This paper analyzes the influence of color dynamic range of input images on the stylized result and proposes two models to improve the spatial rationality of stylized image. Firstly, to solve the unreasonable space caused by the difference in the color dynamic ranges between content image and style image, we propose a neural style transfer model (RNDT) based on robust nonparametric distribution transfer. Then, a neural style transfer model combining exposure corrected and robust nonparametric distribution transfer (EC-RNDT) is proposed to further improve the stylized result of RNDT model for underexposed content images. Experimental results show that the style transfer methods based on robust nonparametric distribution transfer not only increased the spatial rationality, but also extended the application range.

The contributions of the proposed models are listed as follows: (1) We introduce a robust nonparametric distribution transfer algorithm into the style transfer model, which could improve the spatial rationality of the stylized image when the color styles of the content image and style image are quite different; (2) We introduce the adaptive exposure correction algorithm into the robust nonparametric distribution transfer to obtain the satisfactory stylized result for underexposed images.

2. Related Work

Since mid-1990s, different strategies have been used to preserve the structure of content images in the field of non-photorealistic rendering [1,4]. Song et al. [15] and Kolliopoulos [16] used a region-based algorithm to manipulate the geometry so as to control the local details. Hertzmann [17] proposed a stroke-based rendering to guide the virtual strokes placement on a digital canvas to endow the content image a prescribed style. Efros and Freemand [18] constrained the texture placement by the prior map exacted from content image, enabling the preservation of content and structure in content image. Even though the traditional artistic stylization can render a stylized image with original structure, it still lacks efficiency and flexibility.

The first neural style transfer (NST) model was proposed by Gatys et al. [8], which takes advantage of the deep convolutional neural network to extract high-level features and transfers them into arbitrary images. The loss function of this method consists of two parts, which are content loss and style loss. The content loss minimizes the distance of the content image and stylized image in feature space. The style loss is defined based on the difference between two Gram matrices from the style image and stylized image. Although the NST model addressed the limitations of the traditional artistic stylization methods, the details cannot be preserved due to the lack of low-level features.

Since the Gram-matrix based loss has limitations, many researchers found it is not the only choice for style representation. Li et al. [19] proposed a Maximum Mean Discrepancy (MMD) loss, which can be applied in NST. It is proved that optimizing Gram-based loss function equals to minimizing the MMD with a quadratic polynomial kernel. Li and Wand [20] introduced an MRF-based style

loss function in patch level, which could preserve the details. However, it has limitation in depth information. Inspired by MRF-based style loss function [20], Liao et al. [21] proposed a deep image analogy method combining three parts, which are the idea of "image analogy" [22], the features exacted by CNN, and the idea of patch-wise matching. However, it works well only when the content image and style image have similar objects.

In addition, some research focuses on improving the stylized result through different information and strategies. Liu et al. [23] introduced a depth loss function based on the perceptual loss [24] to retain the overall spatial layout. Champandard [25] proposed a semantic-based algorithm by incorporating a semantic channel over the MRF loss [20], which leads to a more accurate semantic match. Li et al. [26] introduced an extra Laplacian loss, which introduces a constraint on low-level features. Gatys et al. [27] and Wang et al. [28] both utilized a coarse-to-fine procedure, enabling a large stroke size and high-resolution image stylization.

Based on the style transfer methods, many existing works focus on a specific application. Luan et al. [29] and Mechrez et al. [30] used a two-stage procedure to perform photorealistic style transfer. Chen et al. [31] proposed a bidirectional disparity loss for stereoscopic style transfer. Ruder et al. [32,33] applied style transfer to video by introducing a temporal consistency loss based on optical flow. Gatys et al. [27] proposed a color-preserving style transfer by performing a luminance-only transfer.

In recent years, the Generative Adversarial Networks (GAN) [34] have been widely used in image-to-image translation, which is similar to the neural style transfer. Isola et al. [35] utilized a conditional GAN [36] to implement the translation of label maps to street scenes, label maps to building facades, object edges to photos. Zhu et al. [37] used a cycle-consistent adversarial networks to translate the photograph and the paintings of famous artists. Image-to-image translation not only transfers the image styles, but also manipulates the attributes of the objects [35,37–40]. Therefore, the image-to-image translation can be seen as a generalization of the neural style transfer [41]. Besides, GANs learns the image styles from multiple images with similar style, while the neural style transfer only needs one content image and one style image.

In this paper, we analyze the factors that affect the rationality of the neural style transfer. Then two style transfer models are proposed based on robust nonparametric distribution transfer. Both style transfer models are verified in different application scenarios. Experimental results show that the proposed models can effectively improve the spatial rationality of the stylized images and preserve the structure information.

The rest of this paper will be organized as follows. We introduce our style transfer models in Section 3. In Section 4, the style transfer results of the proposed models and the existing models are analyzed. Finally, we conclude our work in Section 5.

3. Neural Style Transfer

In this section, we analyze two inevitable factors affecting the space rationality of the style transfer. One is the unreasonable space caused by the difference of the dynamic color ranges, and another one is underexposure. Both factors limit effectiveness of the neural style transfer models. Based on the analysis, we propose two neural style transfer models to ease the limitations and improve the effectiveness of the neural style transfer models. In Section 3.1, the RNDT model is presented to solve the problem of unreasonable space caused by the difference of the dynamic color ranges. In Section 3.2, the EC-RNDT style transfer model introduced the adaptive detail-enhanced multi-scale retinex into the RNDT model, which can simultaneously solve the problems of unreasonable space caused by the difference of the dynamic color ranges, and underexposure. Both methods can be utilized as the preprocessing of other neural style transfer models.

3.1. Robust Nonparametric Distribution Transferred Neural Style Transfer Model

Compared with the traditional artistic stylization, the neural style transfer method can work on any set of content images and style images. However, this approach does not yield a good stylized

result for some input images. Since the colors in style images are usually more saturated than the content images, the color dynamic range of the style image is usually larger than that of the content image. We find that when the color dynamic range of the content image is much smaller than that of the style image, the artistic style texture is prone to being misplaced.

To solve this problem, a neural style transfer model based on robust nonparametric distribution transfer is proposed. Before the style transfer, the color of the style image is transferred into the content image. This method is named robust nonparametric distribution transferred neural style transfer, which is shown in Figure 2.
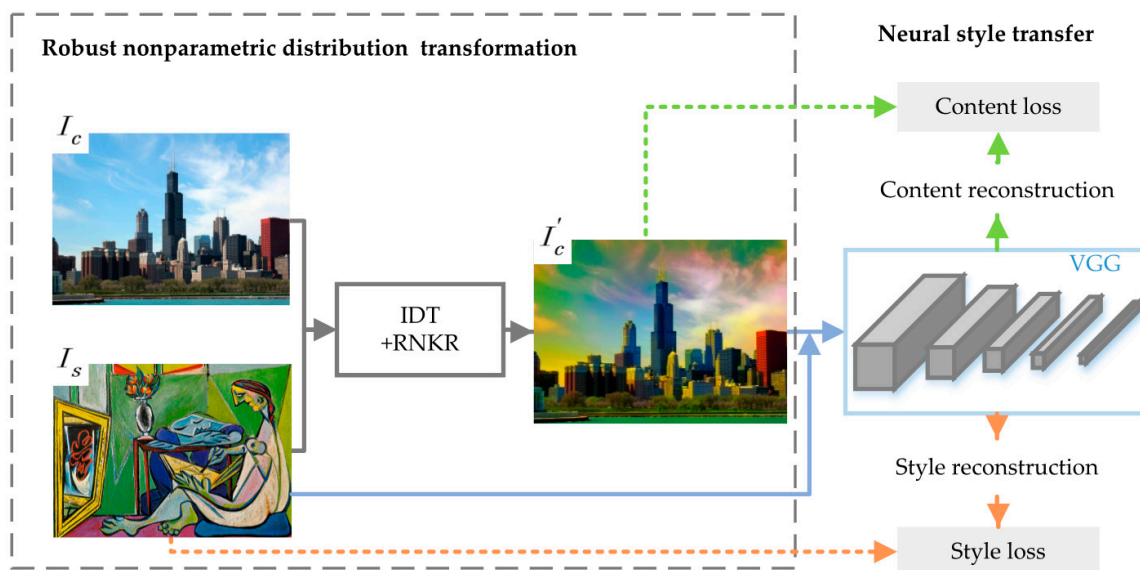

Figure 2. The structure of robust nonparametric distribution transferred neural style transfer.

For the robust nonparametric distribution transfer, it is constructed by introducing the robust nonparametric kernel regression (RNKR) [42] into the iterative Distribution Transform (IDT) model [43,44]. The IDT algorithm can change the color of all pixels into colors in the style image, while protecting the color continuity within the objects. When the color styles of the two input images are very different, and the color dynamic range of each image is very large, the IDT model can still get good results.

The IDT algorithm makes the color probability distribution of content image become that of the style image. Therefore, the color dynamic range of the content image is enhanced. Moreover, the boundaries of each object are more obvious. As a result, the content loss has a larger value at the boundary of the objects. Nonparametric kernel regression is a familiar tool to explore the underlying relation between response variable and covariates. The RNKR model [42] can control outlier effect through combining weighting and trimming to robustly achieve the nonparametric kernel regression. Based on this, RNKR is utilized to replace and obtain the optimal transport solution in IDT, when solving estimating equations in IDT.

The robust nonparametric distribution transfer procedure is presented by Equation (1),

$$I_c' \leftarrow t_s(I_c, I_s) \tag{1}$$

where I_c is the original content image. I_c' is the content image with the color of style image, which will be input into the feature extraction network. $t_s(\cdot)$ is a mapping function, which transforms the color distribution of I_c into that of style image I_s. The input of $t_s(\cdot)$ is the content image I_c and the style image I_s. A new image with the content of I_c and the color of I_s will be generated. The details of the $t_s(\cdot)$ are presented as Algorithm 1 [43,44].

Algorithm 1 The details of the $t_s(\cdot)$

Initialization of the source data I_c and target I_s For example in color transfer, $I_{c(j)} = (r_{(j)}, g_{(j)}, b_{(j)})$ where $r_{(j)}, g_{(j)}, b_{(j)}$ are the red, green, and blue components of pixel j.

$k \leftarrow 0, I_c^{(0)} \leftarrow I_c$

repeat

(1)   take one rotation matrix R

(2)   obtain the rotated samples: $I_{c(r)} \leftarrow RI_c^{(k)}$ and $I_{s(r)} \leftarrow RI_s$

(3)   project the samples on all axes j to get the marginal $f_{(j)}$ and $g_{(j)}$

(4)   for each axis j, find the 1D transformation $t_{s(j)}$ that matches the marginals $f_{(j)}$ and $g_{(j)}$ by using RNKR.

(5)   remap the samples $I_{c(r)}$ according to the 1D transformations by using RNKR. For example, a sample $(I_{c(1)}, \ldots, I_{c(N)})$ is remapped into $(t_1(I_{c(1)}), \ldots, t_N(I_{c(N)}))$, where N is the dimension of the samples.

(6)   rotate back the samples: $I_c^{(K+1)} \leftarrow R^{-1}I_{c(r)}$

(7) $k \leftarrow k + 1$

until convergence on all marginals for every possible rotation

The final one-to-one mapping $t_s(\cdot)$ is given by: $\forall j, I_{c(j)} \rightarrow t_s\left(I_{c(j)}\right) = I_{c(j)}^\infty$

In the neural style transfer model, VGG-19 [6] was used as a feature extractor to extract the high-level feature exaction for the input images and output images. The loss function of the proposed method consists of content loss and style loss. The content loss minimizes the content distance between the content image and the stylized image. The style loss is defined based on the difference between the high-level artistic style from the style image and the stylized image.

When the color dynamic range of the content image is much smaller than that of the style image, the artistic style textures tend to be placed incorrectly. The proposed RNDT model can improve the stylized effect in this case. In practical applications, the stylized result of the RNDT model is more spatially reasonable in this situation.

3.2. Exposure Corrected and Robust Nonparametric Distribution Transferred Neural Style Transfer Model

This section focuses on the daily photo stylization for cellphone users. Sometimes, photos are more likely to be underexposed. The underexposed photos usually have the following problems: (1) in the backlit area, the edge gradient of the object is smaller than normal area. The feature at the edge of the object is weakened so that the content loss function does not work well. (2) The color saturation in the backlit region is low, resulting in a small color dynamic range.

To solve these problems, an adaptive detail-enhanced multi-scale retinex (DEMSR) algorithm for exposure correction is proposed. Based on DEMSR, a neural style transfer model integrating exposure corrected and robust nonparametric distribution transfer is proposed.

3.2.1. Adaptive Detail-Enhanced Multi-Scale Retinex Algorithm

Our proposed exposure correction algorithm for an underexposed content image consists of the following three steps.

Adaptive backlit region detection. The backlit region template detection is used to determine whether to modify the image. The adaptive detection method of backlit region is shown in Equation (2),

$$M(x, y) = n_{0-1}(f(\lambda \times (L_T - L(x, y)))) \tag{2}$$

where L is the lightness channel in CIELab space, L_T is a threshold by which the image pixels are classed as backlit region or non-backlit region, λ is an empirical parameter that represents the degree of data polarization, $f(u)$ is a function mapping the variate from range $(-\infty, +\infty)$ to

range (0,1), $n_{0-1}(z) = (z - \min(Z))/(\max(Z) - \min(Z))$ normalizes the pixels in the image. Here, $f(u) = 0.5/\tanh(1 + u)$, $L_T = 70$, and $\lambda = 60$.

According to [45], the guided filter can protect the edges of large objects in the image while blurring the image. Derived from a local linear model, the guided filter obtains the filtering output by learning the content of the guidance image, which can be the input image itself or another different image. Let F represent a guided filter whose window is 10% of the size of the content image. Then, the final backlit region template M_g can be calculated by Equation (3).

$$M_g = F * M \tag{3}$$

Light restoration module. Our exposure correction algorithm is developed based on the Multi-Scale Retinex (MSR) algorithm [46]. In the MSR algorithm, detail enhancement is performed on the luminance channel Y in YCbCr space. The algorithm estimates the natural light by blurring the luminance channel Y in three scales. Therefore, the reflected light R of the object itself can be estimated by Equation (4), while the exposure corrected luminance Y^{MSR} is formulated by Equation (5).

$$R_i = \log Y - \log L_i, Y(x, y) \in [0, 1] \tag{4}$$

$$Y^{MSR} = n_{0-1}\left(\sum_{i=1}^{3} \frac{1}{3} R_i\right) \tag{5}$$

where $L_i = F_i \times Y, i = 1, 2, 3$, F_i are the Gaussian filters that estimate the natural light in three scales, the window sizes of F_1, F_2, F_3 are 0.8%, 6%, 15% of the content image, respectively. Y^{MSR} is an exposure corrected image in the luminance channel which is calculated by the MSR algorithm.

In order to preserve the fine details after the robust nonparametric distribution transfer of the RNDT model, our method has made two improvements compared to the MSR algorithm. (1) Removing the "Halo" phenomenon to protect the reality of the boundary of the backlit area. For this purpose, the Gaussian filter used in Equation (4) is changed into a guided filter. The guided filter can protect the edge of the backlit region, while estimating the natural light L_i; (2) Enhancing the edge and stereoscopy of the object inside the backlit area. The idea of the Laplacian filter [47] is used to enhance the edge of the object in the enhanced backlit area by Equation (6),

$$Y^{DEMSR} = Y^{MSR} + (Y - F_4 * Y) \tag{6}$$

where F_4 is a Laplacian filter whose window is 0.4% of the content image. Y^{DEMSR} is the exposure corrected image in the luminance channel obtained by the DEMSR algorithm. After enhancing the details, a gain compensation on Y^{DEMSR} will be conducted for the balance of the three channels. Finally, the YCbCr space is converted to the RGB space to obtain the backlit enhanced image I^{DEMSR}.

Fusion to obtain exposure correction image. Once the backlit enhanced image is obtained, the non-backlit region is filtered by the backlit region template. The enhanced backlit area is fused with the normal area to obtain the final exposure correction image I_{ec} by Equation (7).

$$I_{ec} = M_g \cdot I^{DEMSR} + \left(1 - M_g\right) \cdot I_c \tag{7}$$

The final exposure correction image will be input to the RNDT model for getting the stylized image.

3.2.2. EC-RNDT Model

Figure 3 shows the proposed EC-RNDT neural style transfer model combining adaptive DEMSR algorithm and RNDT algorithm. The exposure correction and color distribution transform include three steps. The first step is an automatic detection of whether the image is underexposed. Second, the light

restoration and the enhancement of edge details are performed. Finally, the exposure correction image I_{ec} is input into the RNDT model.

$$I'_{ec} \leftarrow t_s(I_{ec}, I_s), \qquad (8)$$

where I'_{ec} is the exposure correction image with the color of style image, which will be the content image of neural style transfer.
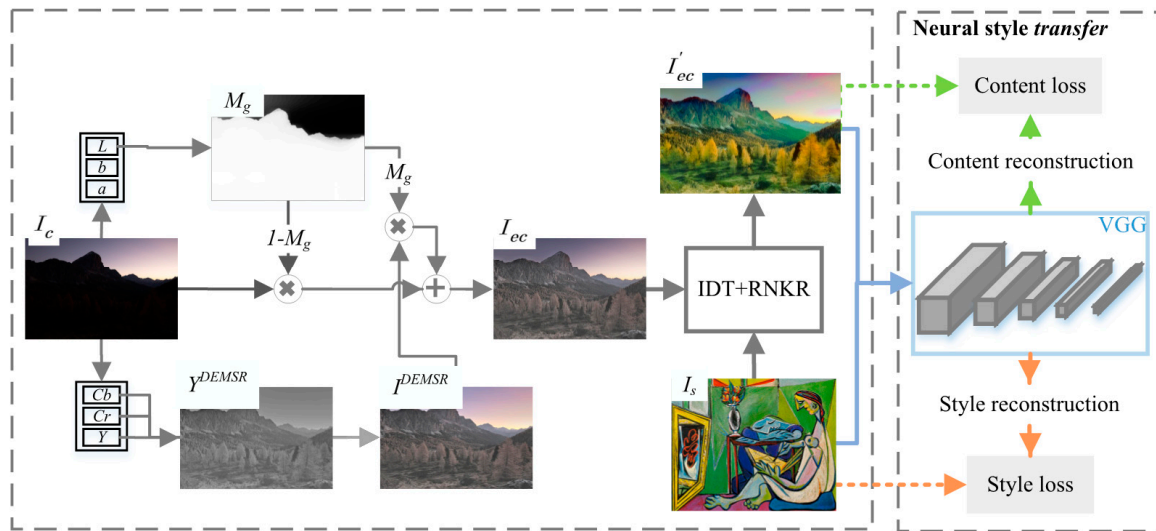

Figure 3. The structure of neural style transfer model using EC-RNDT.

4. Experimental Results

4.1. Experiment Details

The RNDT algorithm and DEMSR algorithm are implemented by Matlab, and the style transfer models use the Lua language and Torch framework. The algorithm was implemented in Python with developed codes subject to the PyTorch framework. The algorithm ran on an Ubuntu system with an Intel Xeon E5-2620 CPU (2.1 GHz) and with a 64 GB memory. Our code used a GTX 1080 Ti GPU with 11 GB memory. The implementation was not optimized and did not use multithread and parallel programming. Computing by using the test images in this section, the average computational cost of RNDT is 1.1 s, and the average computational cost of EC-RNDT is 4 s. In the experiments, all the style transfer models run in the same environment, and all the hyper-parameters are unchanged in the comparison experiments.

It is a subjective task to evaluate the quality of the style transfer [3,4]. Therefore, most of evaluations of neural style transfer models are qualitative [48,49]. The most common method is to qualitatively compare results of style transfer methods by putting stylized images side by side [11–14,21,24,29]. Besides showing stylized images, user study is also used for evaluation [9,14,29]. The typical setup is to recruit some users, show them stylized images of different models, and ask them which results they prefer. The qualitative results will be presented in Sections 4.2–4.5, while the quantitative results of user study are shown in Section 4.6.

4.2. Style Transfer Results of RNDT Model

For the comparison experiments, the models of Gatys et al. [8], Huang et al. [11], Li et al. [9], Liao et al. [21], and Johnson et al. [24] were used as the baselines. The results of these five baseline models and our RNDT model were compared to verify the superiority of the proposed method. The results of Gatys et al. [8] and RNDT model are shown in Figure 4. In the first result of Gatys et al., the color of flowers is similar with the background (yellow box), resulting in an ambiguous edge of the flower. It also has ambiguous edge for the sail in the second row (yellow box). However, after the

robust nonparametric distribution transfer of content image, the proposed model can generate more reasonable stylized results. The distribution of color is more reasonable, and the boundaries of the objects are much clearer than the left result.
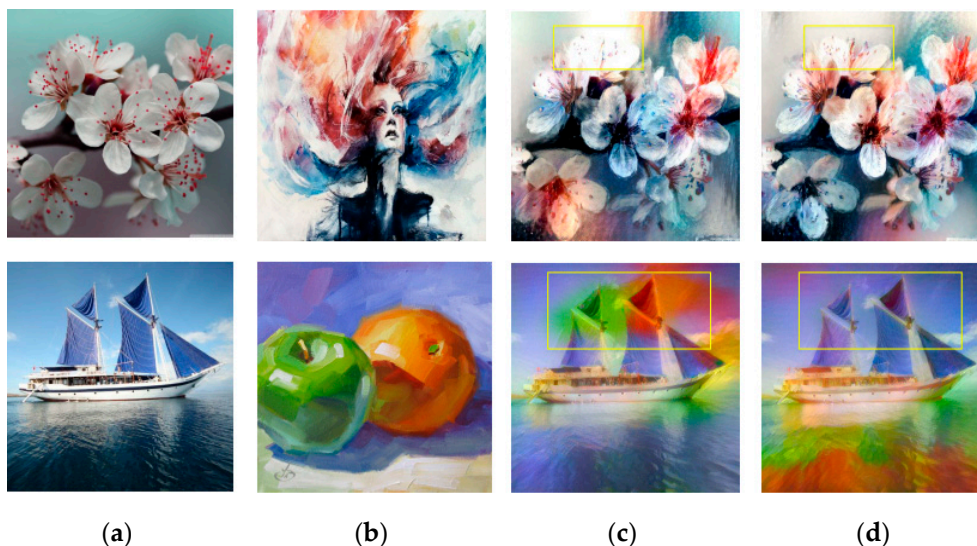

(a)   (b)   (c)   (d)

Figure 4. The style transfer results of Gatys et al. and RNDT model: (**a**) content image; (**b**) style image; (**c**) Gatys et al.; (**d**) RNDT.

Figure 5 shows the stylized results of Huang et al. [11] and RNDT model. In the first row, the model of Huang et al. does not transfer the primary color of the buildings in the style image into the content image. At the same time, the model does not transfer the red color in a proper location in the second row, and the spatial distribution of colors in the result is unreasonable. The proposed robust nonparametric distribution transfer solves these problems. In the first row, the RNDT model transfers the yellow color in the style image to the output, and preserves the structures. In the second result of RNDT model, the colors inside the sails are more continuous. The stylized results of our method have clearer edges than model of Huang. Similarly, other three comparison results are presented in Figures 6–8.
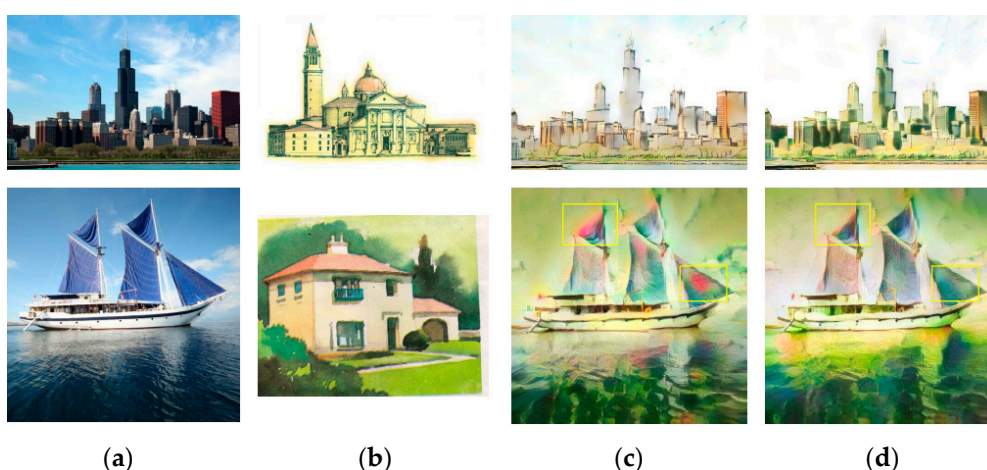

(a)   (b)   (c)   (d)

Figure 5. The style transfer results of Huang et al. and RNDT model: (**a**) content image; (**b**) style image; (**c**) Huang et al.; (**d**) RNDT.
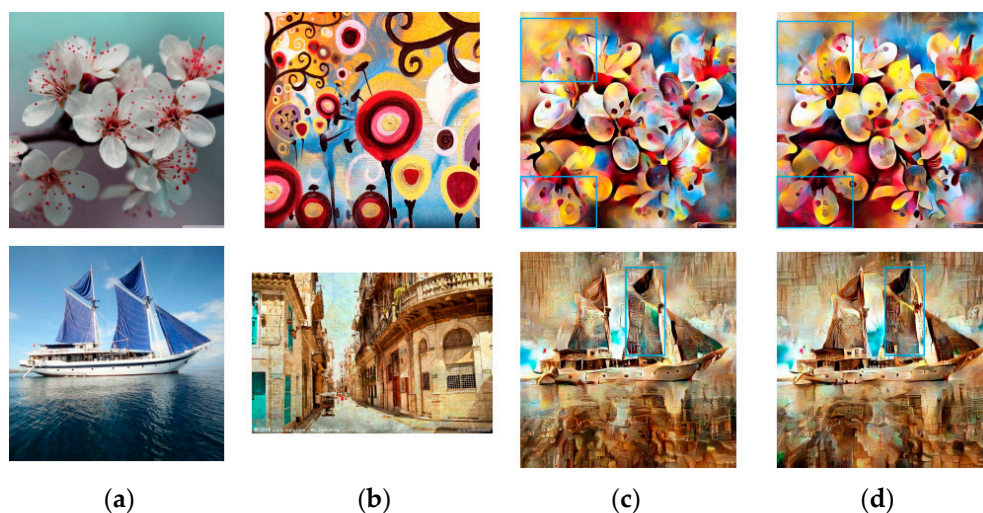
Figure 6. The style transfer results of Li et al. and RNDT model: (**a**) content image; (**b**) style image; (**c**) Li et al.; (**d**) RNDT.
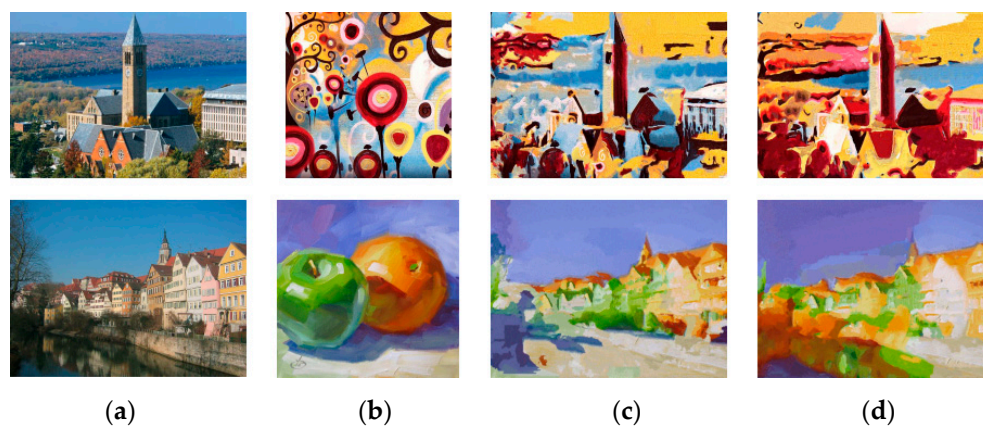

Figure 7. The style transfer results of Liao et al. and RNDT model: (**a**) content image; (**b**) style image; (**c**) Liao et al.; (**d**) RNDT.
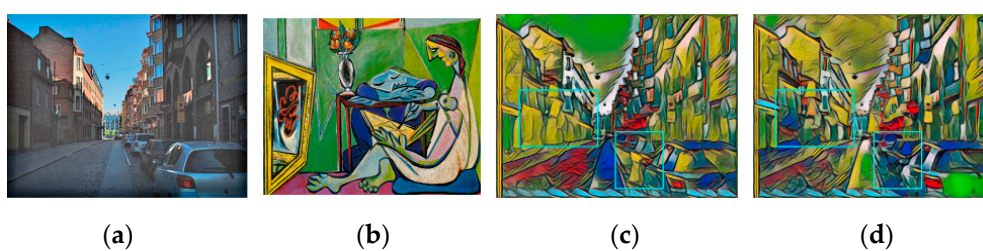

Figure 8. The style transfer results of Johnson et al. and RNDT model: (**a**) content image; (**b**) style image; (**c**) Johnson et al.; (**d**) RNDT.

When the color dynamic range of the content image is much smaller than that of the style image, the artistic style texture is prone to be misplaced. An experiment was performed to show that the existing neural style transfer models suffer from this problem. The results are shown in Figure 9. The first row shows the results where the color dynamic range of the content image is smaller than that of the style image. We can see that the texture of stylized image is misplaced (yellow box). Meanwhile, when the color dynamic ranges of the content image and style image are close, as shown in the second row of the figure, the texture is placed correctly.
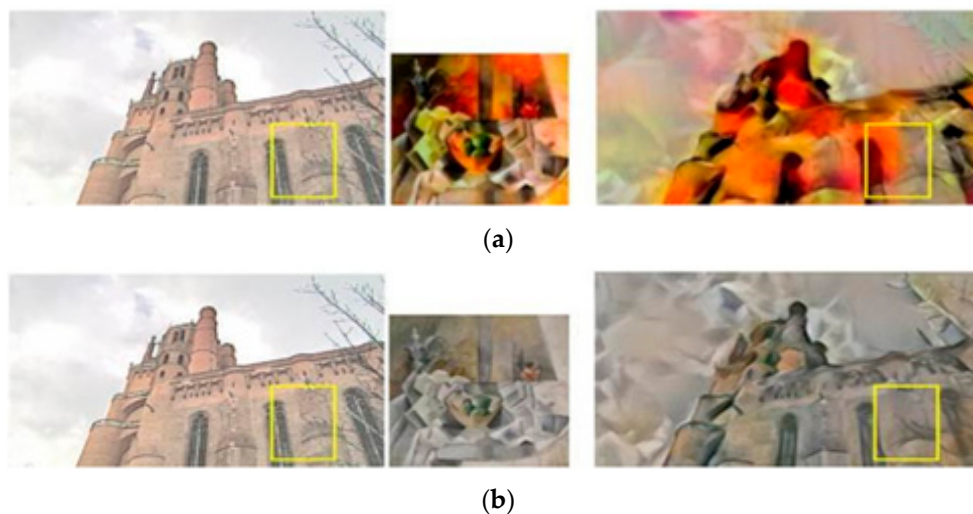
(a)

(b)

Figure 9. The impact of different color dynamic ranges between the content image and style image on the results: (**a**) when the color dynamic range of I_c is smaller than I_s, the texture is misplaced; (**b**) when the color dynamic range of I_c is close to I_s, the texture is placed correctly.

To evaluate the impact of the RNDT algorithm, an experiment was performed. The experiment results are shown in Figure 10, where I'_c is the content image with the color of style image. When the color dynamic range of content image is far smaller than that of style image, as shown in Figure 10a,b, the method of Gatys et al. [8] cannot maintain the spatial structure of the content image. As shown in the yellow box of Figure 10e, sky area and building area are wrongly mixed together, while the RNDT method could overcome this defect. As shown in the yellow box of Figure 10f,g, the sky and building are distinguished, which benefit from the color pre-transformation. Furthermore, we find that the result based on RNDT is better than that of using Reinhard [50]. As shown in the blue box of Figure 10f,g, the result based on RNDT has more uniform color inside the building area. It could be explained based on the intermediate results of color pre-transformation, which are shown in Figure 10c,d. The linear color transformation algorithm of Reinhard tends to produce an excessive discoloration. However, the nonlinear RNDT algorithm has no excessive color change, which could maintain the spatial structure of the object.
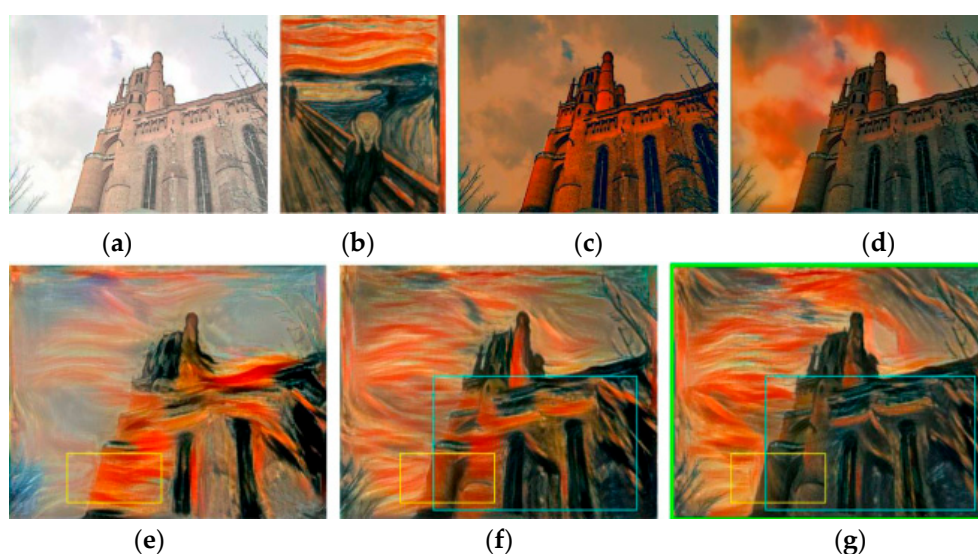

(a)　　　　　(b)　　　　　(c)　　　　　(d)

(e)　　　　　(f)　　　　　(g)

Figure 10. Comparison results: (**a**) content image; (**b**) style image; (**c**) I'_c based on Reihard; (**d**) I'_c based on RNDT; (**e**) transfer result of Gatys; (**f**) result based on Reihard; (**g**) result based onRNDT.

4.3. Results of DEMSR Algorithm

　　To illustrate the effectiveness of the DEMSR algorithm, the intermediate results are shown in Figure 11. These results correspond to different exposure correction steps in DEMSR, respectively. More specifically, Figure 6a shows an underexposed content image. Figure 11b,c shows the outputs of the adaptive backlit region detection. Figure 11d,e shows the results of the light restoration module. Finally, Figure 11 is the exposure correction image of the proposed DEMSR algorithm. Through the DEMSR algorithm, the light was restored and the details in the backlit region were enhanced.
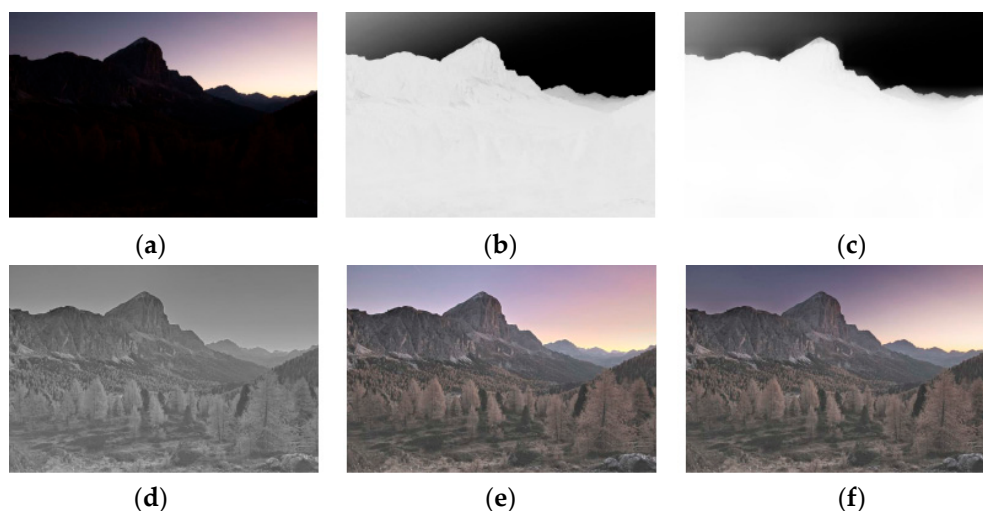

Figure 11. The illustration of the Adaptive Detail-Enhanced Multi-Scale Retinex algorithm: (**a**) content image; (**b**) initial backlit region template; (**c**) final backlit region template; (**d**) Y channel of backlit enhanced image; (**e**) backlit enhanced image; (**f**) exposure correction image.

　　The light restoration results of the MSR algorithm and our DEMSR algorithm are compared in Figure 12. Since the object edge inside the backlit region was enhanced in the DEMSR, the output of DEMSR can show the details more clearly than the MSR algorithm.
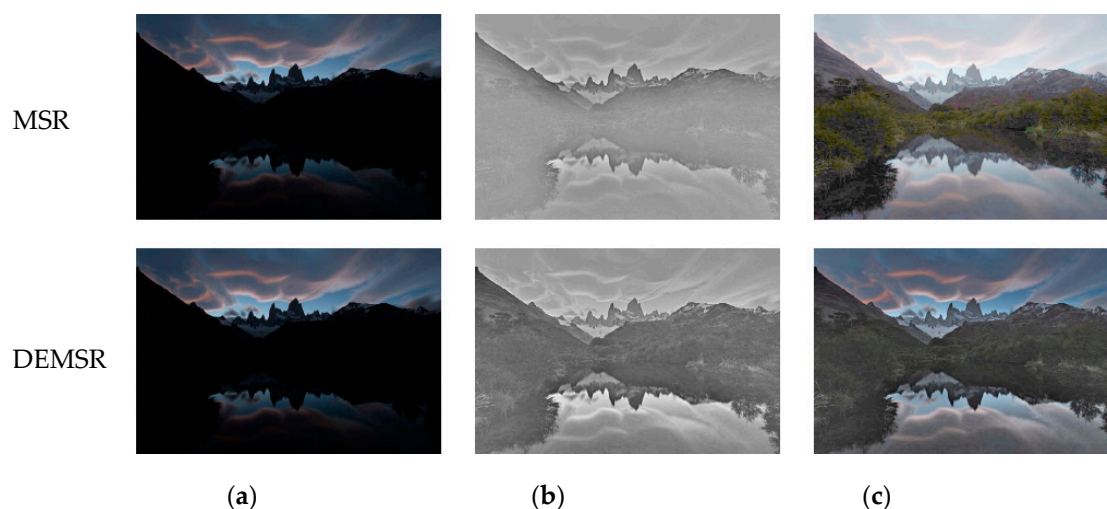

Figure 12. The comparison between MSR and the proposed DEMSR algorithms: (**a**) source image; (**b**) Y channel of backlit enhanced image; (**c**) exposure correction image.

　　To further evaluate the performance of MSR and the proposed DEMSR algorithms, two quantitative metrics are adopted, which are peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [51]. The bigger values of these two metrics mean better performance. The PSNR obtained from MSR

algorithm is 8.35, while the PSNR obtained from DEMSR is 11.85. The MSR algorithm gets a SSIM of 0.26, while the proposed DEMSR gets a SSIM of 0.39. Therefore, the proposed DEMSR algorithm performs better than the MSR algorithm.

The adjustment of parameters L_T and λ in the backlit region detection determines the effect of exposure correction. Therefore, we performed an experiment to evaluate the influences of these two parameters. The results obtained based on different values of the parameters are shown in Figure 13. From the figure, we can see that there is only a slight change between different style transfer results. Therefore, the proposed method is robust to the parameters L_T and λ in a fixed range, e.g., L_T varies from 60 to 80, while λ varies from 50 to 70.
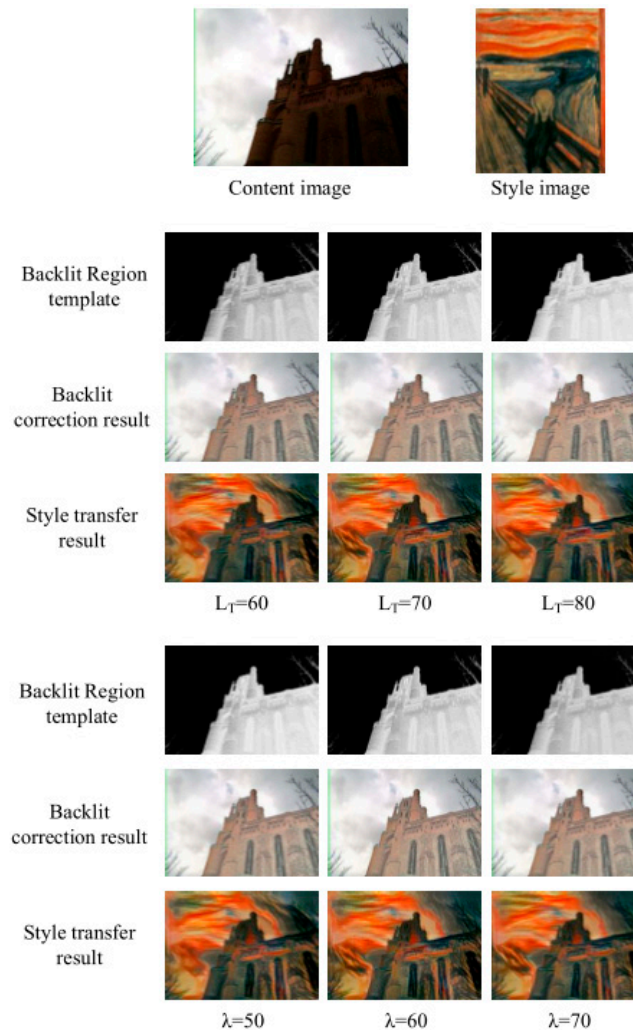

Figure 13. The influences of the backlit-region-detection parameters L_T and λ.

4.4. Style Transfer Result of EC-RNDT Model

4.4.1. Comparison of RNDT Model and EC-RNDT Model

In order to show the effect of the DEMSR algorithm, we compared the results of the RNDT model with the EC-RNDT model.

In Figure 14, the output image of Gatys et al. [8], RNDT, and EC-RNDT model on the underexposed content image are shown, respectively. Figure 8a is the underexposed content image and Figure 8b is the style image. Figure 8c is the result of robust nonparametric distribution transfer in the RNDT model. In Figure 8d, M_g is the backlit template, I_{ec} is the result of DEMSR algorithm, and I'_{ec} is the

result of the robust nonparametric distribution transfer in the EC-RNDT model. Figure 8e–g shows the stylized results of Gatys et al., RNDT model, and EC-RNDT model, respectively.
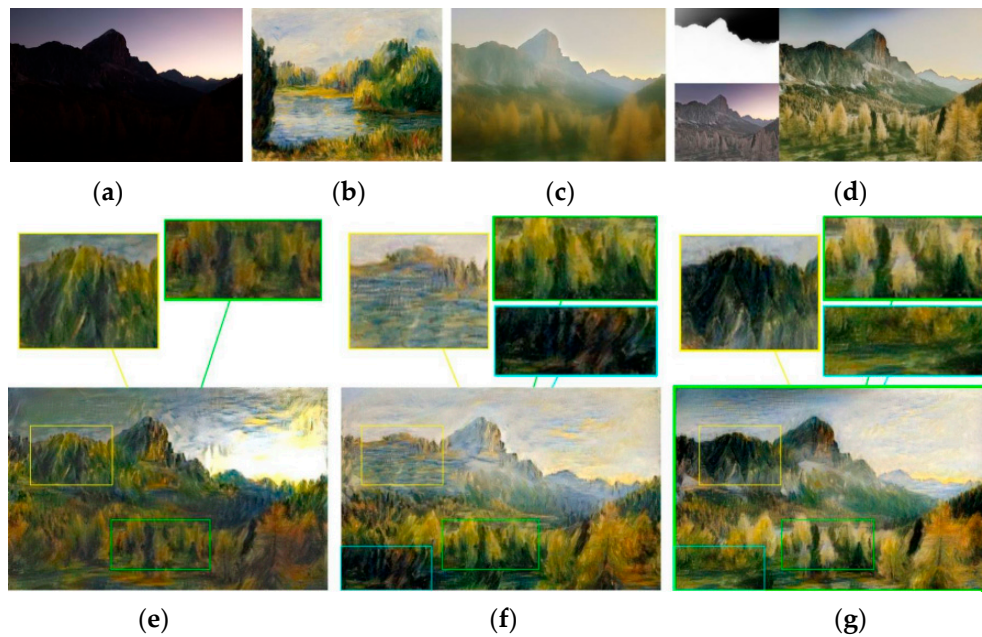

Figure 14. The results of Gatys et al., RNDT and EC-RNDT model on the underexposed content image: (**a**) content image; (**b**) style image; (**c**) I'_c; (**d**) M_g, I_{ec}, I'_{ec}; (**e**) Gatys et al.; (**f**) RNDT; (**g**) EC-RNDT.

In Figure 14, the tree and hill in the stylized result of Gatys et al. have the same color and texture (yellow and green boxes), which makes the structures unclear. In the stylized result of RNDT, the texture of the tree and hill become different due to the robust nonparametric distribution transfer. However, the details in the stylized image are still unclear, such as the boundary between grasses and trees (the blue box), the details inside the hill (the yellow box), and the boundary between different trees (the green box). The EC-RNDT model further improves the stylized result by using the exposure correction. Since the DEMSR algorithm enhances the details in the backlit region, the EC-RNDT model can preserve the fine details in the content image. Benefiting from the DEMSR algorithm, the EC-RNDT model can better protect the spatial rationality of the stylized image than RNDT model when the content image is underexposed.

4.4.2. Comparison between Baselines and EC-RNDT Mode

In order to show the superiority of the proposed EC-RNDT model, we compared four baseline models [8,11,21,24] with the EC-RNDT model.

In the experiments, the models of Gatys et al. [8] and Huang et al. [11] were used as the baselines. The stylized results of baselines [8,11] and EC-RNDT model on an underexposed content image were compared. These results show the effectiveness of the exposure correction and robust nonparametric distribution transfer.

The stylized results of Gatys et al. [8] and EC-RNDT model are shown in Figure 15. As shown in the green box, the stylized results of Gatys et al. loose some details in the building, resulting in an ambiguous structure of the building. Comparing Figure 15c,d, it is easily found that the EC-RNDT model can render the color and preserve the structure better. It can be attributed to exposure correction preforming before the style transfer, which can learn and preserve the fine details in the backlit region well.

The stylized results of Liao et al. and EC-RNDT model are shown in Figure 16. From the first row, it is obvious that the structures of the building obtained by EC-RNDT are much clearer than that of

Huang et al. Meanwhile, the EC-RNDT model renders more colorful results while Huang et al. cannot effectively transfer all the colors of the style image. In the second row, the trees and the reflection in the stylized image of Huang et al. have similar color and texture, which makes the image content blurry. For EC-RNDT model, the different kinds of object have different colors and textures. Therefore, the EC-RNDT model can effectively retain the image content, and transfer the texture and color styles. Similar comparison results are presented in Figures 17 and 18.
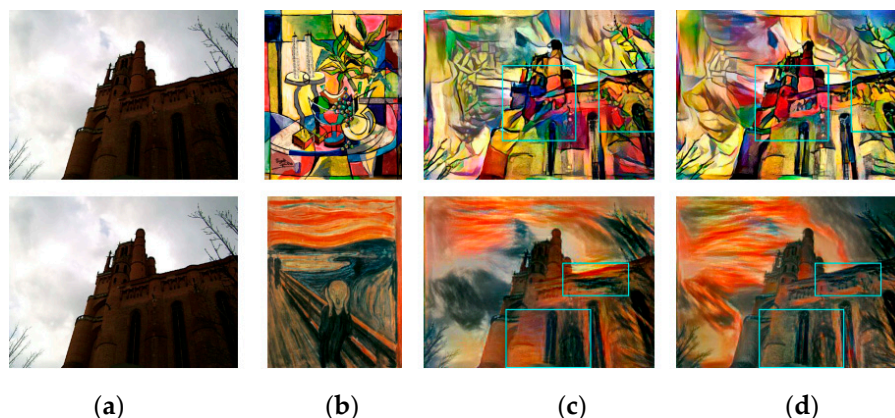

(**a**)   (**b**)   (**c**)   (**d**)

Figure 15. The style transfer results of Gatys et al. and EC-RNDT model: (**a**) content image; (**b**) style image; (**c**) Gatys et al.; (**d**) EC-RNDT.
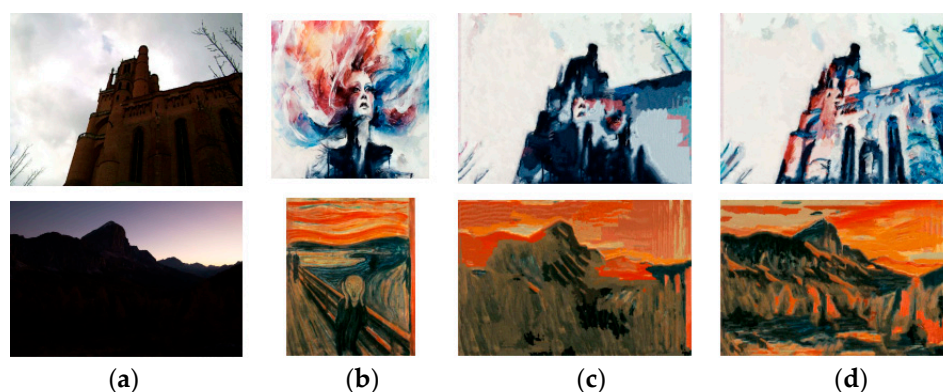

(**a**)   (**b**)   (**c**)   (**d**)

Figure 16. The style transfer results of Liao et al. and EC-RNDT model: (**a**) content image; (**b**) style image; (**c**) Liao et al.; (**d**) EC-RNDT.

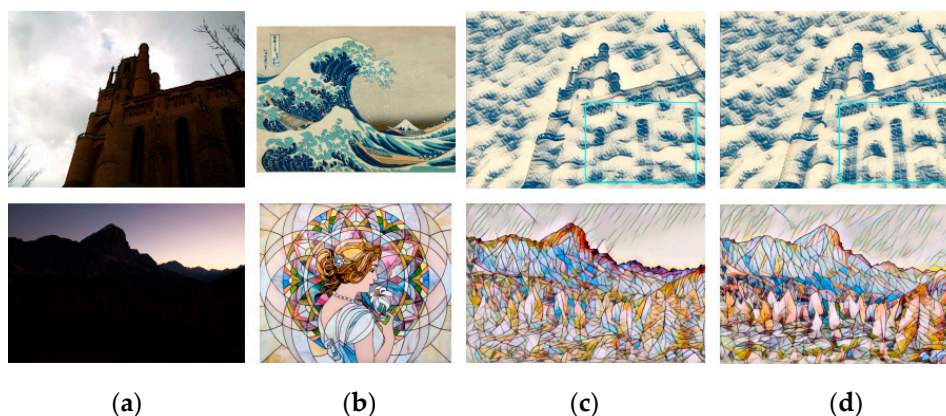(**a**)   (**b**)   (**c**)   (**d**)

Figure 17. The style transfer results of Johnson et al. and EC-RNDT model: (**a**) content image; (**b**) style image; (**c**) Johnson et al.; (**d**) EC-RNDT.
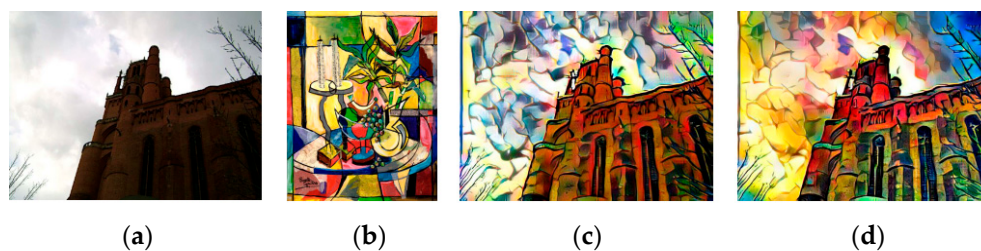
Figure 18. The style transfer results of Huang et al. and EC-RNDT model: (**a**) content image; (**b**) style image; (**c**) Huang et al.; (**d**) EC-RNDT.

4.5. Color-Preserved Style Transfer

Aforementioned sections showed the color transfer of the content image. Under certain circumstances, style transfer is hoped to preserve the color of the content image but transfer the texture style. In this case, the RNDT model can achieve this demand through a slight modification. The input image of RNDT model performs a role exchange, which can realize the color-preserved transfer. That is to say, before the style transfer model, the color of style image is changed into the color of the content image. Therefore, the style transfer will only transfer the texture style of the style image. The results of color-preserved RNDT model are shown in Figure 19.
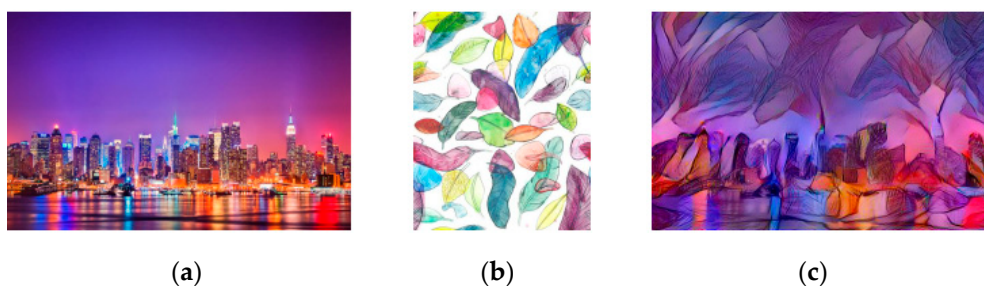

Figure 19. The result of color-preserved RNDT model: (**a**) content image; (**b**) style image; (**c**) color-preserved result.

4.6. Quantitative Results

A user study was performed to further valid the superiority of the proposed models. In the user study, we designed a webpage including 60 groups of outputs, which were viewed by participants. Each group contains two stylized image of baseline models and one stylized image from the proposed model. Participants were asked to choose the one he or she likes better. The results of user study are shown in Tables 1 and 2. The values in the table presents the preference (the percentage of the votes) on the stylized results of different models. It shows that the stylized results of proposed models are better than the stylized results from baseline models. From the user study, we conclude that the proposed models can render more artistic and popular results.

As a form of preprocessing, the proposed model can improve the performances of the corresponding style transfer method without depending on any specific neural style transfer model. To evaluate this, two quantitative metrics were adopted, which are deception rate [52] and FID core [53]. As defined in [52], deception rate is proposed for an automatic evaluation of style transfer results. It is calculated as the fraction of generated images which were classified by the network as the artworks of an artist for which the stylization was produced. FID score evaluates the style transfer results by measuring the distance between the generated distribution and the real distribution [53]. The higher deception rate means the better performance, while the lower FID score represents the better performance. Gatys [8] and Huang [11] were used as the compared methods. Then the proposed model was used as preprocessing of Gatys [8] and Huang [11]. The deception rate and FID core of

different methods are shown in Tables 3–6 separately. From the tables, we can see that the proposed method outperforms two state-of-the-art methods.

Table 1. The preference (%) of stylized images of baseline models and RNDT model when the color dynamic ranges of two input images are much different.

Model	Gatys et al.	Huang et al.	RNDT
Preference	33.60	24.11	42.49

Table 2. The preference (%) of stylized images of baseline models and EC-RNDT model on the underexposed content images.

Model	Gatys et al.	Huang et al.	EC-RNDT
Preference	31.21	20.95	47.84

Table 3. Deception rate on style transfer in terms of different methods. Higher values indicate better performance.

Model	Gatys [8]	Ours + Gatys
Deception rate	0.43	0.49

Table 4. Deception rate on style transfer in terms of different methods. Higher values indicate better performance.

Model	Huang [11]	Ours + Huang
Deception rate	0.31	0.32

Table 5. FID score on style transfer in terms of different methods. Lower score indicates better performance.

Model	Gatys [8]	Ours + Gatys
FID score	265.3	262.8

Table 6. FID score on style transfer in terms of different methods. Lower score indicates better performance.

Model	Huang [11]	Ours + Huang
FID score	245.8	241.7

5. Conclusions

In this paper, we propose two models for image neural style transfer. The first model named as RNDT transfers the color of the content image to that of the style image, which makes the spatial structure of the stylized image more reasonable. The second model introduces the DEMSR exposure correction algorithm to enhance the details in the underexposed images, which can preserve the fine details in the stylized images. Experimental results show that the models have improved the spatial rationality and the artistic sense of the stylized images. The proposed models could be applied in several situations, which include underexposed images, large color-range images, and color-preserved style transfer. Therefore, our models have a high application value.

The proposed RNDT and EC-RNDT models do not depend on any specific neural style transfer model. They work for the style transfer methods that only need one style image. They can improve the performances of the corresponding style transfer methods as a method of preprocessing.

The proposed models achieve encouraging performance in generating stylized images. However, we have to note that our approach still has limitation. The proposed models are not end-to-end, which is not efficient. Therefore, developing an efficient end-to-end style transfer method is our future work.

Author Contributions: Conceptualization, S.L.; methodology, S.L.; software, S.L.; validation, S.L. and Z.T.; formal analysis, C.H.; writing—original draft preparation, S.L. and J.H. All authors have read and agreed to the published version of the manuscript.

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

References

1. Kyprianidis, J.E.; Collomosse, J.; Wang, T.; Isenberg, T. State of the "art": A taxonomy of artistic stylization techniques for images and video. *IEEE Trans. Vis. Comput. Graph.* **2013**, *19*, 866–885. [CrossRef] [PubMed]
2. Semmo, A.; Isenberg, T.; Döllner, J. Neural style transfer: A paradigm shift for image-based artistic rendering? In Proceedings of the Symposium on Non-Photorealistic Animation and Rendering, Los Angeles, CA, USA, 29–30 July 2017; p. 5.
3. Jing, Y.; Yang, Y.; Feng, Z.; Ye, J.; Yu, Y.; Song, M. Neural style transfer: A review. *IEEE Trans. Vis. Comput. Graph.* **2019**. [CrossRef]
4. Rosin, P.; Collomosse, J. *Image and Video-Based Artistic Stylization*; Springer: Berlin/Heidelberg, Germany, 2012.
5. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the 26th Annual Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
6. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2016**, arXiv:1409.1556.
7. Le Cun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [CrossRef] [PubMed]
8. Gatys, L.A.; Ecker, A.S.; Bethge, M. Image style transfer using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2414–2423.
9. Li, Y.; Fang, C.; Yang, J.; Wang, Z.; Lu, X.; Yang, M.H. Universal style transfer via feature transforms. *arXiv* **2017**, arXiv:1705.08086.
10. Shen, F.; Yan, S.; Zeng, G. Neural style transfer via meta networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8061–8069.
11. Huang, X.; Belongie, S. Arbitrary style transfer in real-time with adaptive instance normalization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1501–1510.
12. Karayev, S.; Trentacoste, M.; Han, H.; Agarwala, A.; Darrell, T.; Hertzmann, A.; Winnemoeller, H. Recognizing image style. *arXiv* **2013**, arXiv:1311.3715.
13. Yoo, J.; Uh, Y.; Chun, S.; Kang, B.; Ha, J.W. Photorealistic style transfer via wavelet transforms. *arXiv* **2019**, arXiv:1903.09760.
14. Li, X.; Liu, S.; Kautz, J.; Yang, M.H. Learning linear transformations for fast image and video style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–21 June 2019; pp. 3809–3817.
15. Song, Y.Z.; Rosin, P.L.; Hall, P.M.; Collomosse, J.P. Arty shapes. In *Computational Aesthetics*; The Eurographics Association: Aire-la-Ville, Switzerland, 2008; pp. 65–72.
16. Kolliopoulos, A. *Image Segmentation for Stylized Non-Photorealistic Rendering and Animation*; University of Toronto: Toronto, ON, Canada, 2005.
17. Hertzmann, A. Painterly rendering with curved brush strokes of multiplesizes. In Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, Orlando, FL, USA, 19–24 July 1998; pp. 453–460.

18. Efros, A.A.; Freeman, W.T. Image quilting for texture synthesis and transfer. In Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles, CA, USA, 12–17 August 2001; pp. 341–346.

19. Li, Y.; Wang, N.; Liu, J.; Hou, X. Demystifying neural style transfer. *arXiv* **2017**, arXiv:1701.01036.

20. Li, C.; Wand, M. Combining markov random fields and convolutional neural networks for image synthesis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2479–2486.

21. Liao, J.; Yao, Y.; Yuan, L.; Hua, G.; Kang, S.B. Visual attribute transfer through deep image analogy. *arXiv* **2017**, arXiv:1705.01088. [CrossRef]

22. Hertzmann, A.; Jacobs, C.E.; Oliver, N.; Curless, B.; Salesin, D.H. Image analogies. In Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles, CA, USA, 12–17 August 2001; pp. 327–340.

23. Liu, X.C.; Cheng, M.M.; Lai, Y.K.; Rosin, P.L. Depth-aware neural style transfer. In Proceedings of the Symposium on Non-Photorealistic Animation and Rendering, Los Angeles, CA, USA, 29–30 July 2017; p. 4.

24. Johnson, J.; Alahi, A.; Li, F.-F. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 694–711.

25. Champandard, A.J. Semantic style transfer and turning two-bit doodles into fine artworks. *arXiv* **2016**, arXiv:1603.01768.

26. Li, S.; Xu, X.; Nie, L.; Chua, T.S. Laplacian-steered neural style transfer. In Proceedings of the 25th ACM International Conference on Multimedia, Mountain View, CA, USA, 23–27 October 2017; pp. 1716–1724.

27. Gatys, L.A.; Ecker, A.S.; Bethge, M.; Hertzmann, A.; Shechtman, E. Controlling perceptual factors in neural style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3985–3993.

28. Wang, X.; Oxholm, G.; Zhang, D.; Wang, Y.F. Multimodal transfer: A hierarchical deep convolutional neural network for fast artistic style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5239–5247.

29. Luan, F.; Paris, S.; Shechtman, E.; Bala, K. Deep photo style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4990–4998.

30. Mechrez, R.; Shechtman, E.; Zelnik-Manor, L. Photorealistic style transfer with screened poisson equation. *arXiv* **2017**, arXiv:1709.09828.

31. Chen, D.; Yuan, L.; Liao, J.; Yu, N.; Hua, G. Stereoscopic neural style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6654–6663.

32. Ruder, M.; Dosovitskiy, A.; Brox, T. Artistic style transfer for videos. In *German Conference on Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 26–36.

33. Ruder, M.; Dosovitskiy, A.; Brox, T. Artistic style transfer for videos and spherical images. *Int. J. Comput. Vis.* **2018**, *126*, 1199–1219.

34. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, S.; Bengio, Y. Generative adversarial nets. In Proceedings of the Annual Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.

35. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.

36. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.

37. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.

38. Choi, Y.; Choi, M.; Kim, M.; Ha, J.M.; Kim, S.; Choo, J. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8789–8797.

39. Chang, H.; Lu, J.; Yu, F.; Finkelstein, A. Pairedcyclegan: Asymmetric style transfer for applying and removing makeup. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 40–48.

40. Yi, Z.; Zhang, H.; Tan, P.; Gong, M. Dualgan: Unsupervised dual learning for image-to-image translation. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2849–2857.

41. Liu, M.Y.; Huang, X.; Mallya, A.; Karras, T.; Aila, T.; Lehtinen, J.; Kautz, J. Few-shot unsupervised image-to-image translation. *arXiv* **2019**, arXiv:1905.01723.

42. Huang, H.; Yu, P.S.; Wang, C. An introduction to image synthesis with generative adversarial nets. *arXiv* **2018**, arXiv:1803.04469.

43. Pitié, F.; Kokaram, A.C.; Dahyot, R. Automated colour grading using colour distribution transfer. *Comput. Vis. Image. Underst.* **2007**, *107*, 123–137. [CrossRef]

44. Pitié, F.; Kokaram, A.C.; Dahyot, R. N-dimensional probability density function transfer and its application to color transfer. In Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV'05), Beijing, China, 24–28 October 2010; IEEE: Piscataway, NJ, USA, 2005; pp. 1434–1439.

45. He, K.; Sun, J.; Tang, X. Guided image filtering. *IEEE Trans. Patern. Anal.* **2012**, *35*, 1397–1409. [CrossRef]

46. Cho, S.; Shrestha, B.; Joo, H.J.; Hong, B. Improvement of retinex algorithm for backlight image efficiency. In *Computer Science and Convergence*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 579–587.

47. Kong, H.; Akakin, H.C.; Sarma, S.E. A generalized laplacian of gaussian filter for blob detection and its applications. *IEEE Trans. Cybern.* **2013**, *43*, 1719–1733. [CrossRef]

48. Mould, D. Authorial subjective evaluation of non-photorealistic images. In Proceedings of the Workshop on Non-Photorealistic Animation and Rendering, Vancouver, BC, Canada, 8–10 August 2014; pp. 49–56.

49. Isenberg, T.; Neumann, P.; Carpendale, S.; Sousa, M.C.; Jorge, J.A. Non-photorealistic rendering in context: An observational study. In Proceedings of the 4th International Symposium on Non-Photorealistic Animation and Rendering, Annecy, France, 5–7 June 2006; pp. 115–126.

50. Reinhard, E.; Adhikhmin, M.; Gooch, B.; Shirley, P. Color transfer between images. *IEEE Comput. Graph. Appl.* **2001**, *21*, 34–41. [CrossRef]

51. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]

52. Sanakoyeu, A.; Kotovenko, D.; Lang, S.; Ommer, B. A style-aware content loss for real-time hd style transfer. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 698–714.

53. Chen, X.; Xu, C.; Yang, X.; Song, L.; Tao, D. Gated-gan: Adversarial gated networks for multi-collection style transfer. *IEEE Trans. Image Process.* **2018**, *28*, 546–560. [CrossRef] [PubMed]