Deep Restoration of Vintage Photographs From Scanned Halftone Prints

Qifan Gao
Shanghai Jiao Tong University
gaoqifan@sjtu.edu.cn

Xiao Shu
McMaster University
shux@mcmaster.ca

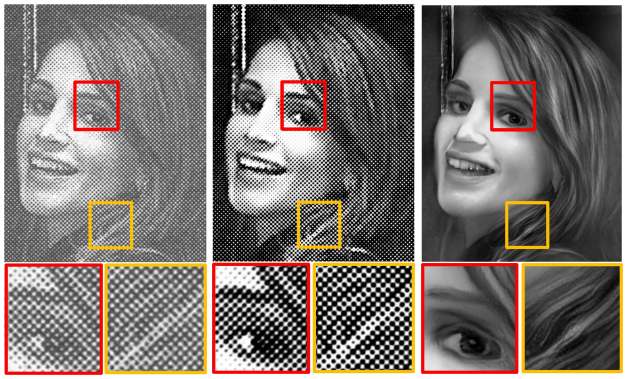
Xiaolin Wu†
McMaster University
xwu@ece.mcmaster.ca

Abstract

A great number of invaluable historical photographs unfortunately only exist in the form of halftone prints in old publications such as newspapers or books. Their original continuous-tone films have long been lost or irreparably damaged. There have been attempts to digitally restore these vintage halftone prints to the original film quality or higher. However, even using powerful deep convolutional neural networks, it is still difficult to obtain satisfactory results. The main challenge is that the degradation process is complex and compounded while almost no paired real data is available for machine learning. In this research, we develop a novel learning strategy, in which the restoration task is divided into two stages: the removal of printing artifacts and the inverse of halftoning. The advantage of our technique is that only the simple first stage, which makes the method adapt to real halftone prints, requires unsupervised training, while the more complex second stage of inverse halftoning only uses synthetic training data. Extensive experiments demonstrate the efficacy of the proposed technique for real halftone prints; the new technique significantly outperforms the existing ones in visual quality.

1. Introduction

In late 1830s camera photographs were first successfully developed on paper sensitised with silver salt [22]. To mass reproduce these photographs on regular paper with black ink, William Henry Fox Talbot invented halftone printing, a technique that uses ink dots of different sizes to simulate different greyscales, shortly after the birth of photography [38]. The first examples of printed photos using this technique appeared in an American periodical as early as in 1873. Since then, countless of valuable historical photographs have been preserved in the form of halftone prints, while their originals were irreversibly damaged or lost. Today, the only possibility to reproduce these lost original photographs is the process of inverse halftoning.



(a) Original (b) Cleaned (c) Restored

Figure 1: The proposed technique divides the task of real halftone print restoration into two stages: the removal of printing artifacts (a)→(b) and the inverse of halftoning (b)→(c). Only the relatively simple first stage requires unsupervised training. The second stage of inverse halftoning can be easily trained with synthetic data.

Recently much publicized successes of deep learning in image restoration stimulated interests in the use of machine learning methods for inverse halftoning. Several deep convolutional neural network (DCNN) based techniques have been proposed to tackle the problem [14, 42, 41, 19]. Their general idea is to train an end-to-end mapping using fabricated image pairs (Y, X), where Y is the digital halftoning result of the latent continuous-tone photograph X. This practice is, however, a seriously flawed expediency. None of the DCNN methods can satisfactorily invert scanned halftone prints of historical photographs, because the dot patterns of relief halftone printing process from the 1880s to the 1950s are very different from those of digital halftoning. A notorious weakness of the black box DCNN methods for image restoration is that even very small deviations between the images for training and inference can cause them to fail. To aggravate the problem further, scanned old halftone prints have other blemishes as well, such as ink smear, pa-

This research is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

†The work was done during this author's visit to Shanghai Jiao Tong University.

per fiber structures, paper damage, etc. How to overcome these difficulties in learning-based inverse halftoning is the technical focus of this research.

Deviating from a direct end-to-end DCNN mapping, the main stream deep learning approach in image restoration, we adopt a new strategy of two-stage deep learning as exemplified in Figure 1. The two stages are two cascaded DCNNs, N_s and N_i, for two tightly coupled learning tasks. The back end subnet N_i is for the task of inverse halftoning, i.e., the inference of X, but the input of N_i is synthetic halftone images $\hat{Y}$ rather than real scanned halftone prints Y. The synthetic halftone images $\hat{Y}$ play the role of circumventing the obstacle of no paired vintage photos and their halftone prints for training. These intermediate training halftone images $\hat{Y}$ are generated by a physically-based model M, i.e., $\hat{Y} = M(X)$.

However, this realistic model M cannot simulate all the compounded noise and artifacts of a vintage halftone print; its outputs $\hat{Y}$ are still only approximations of the pertaining real halftone prints Y that deep learning desires but cannot have. The model precision gap is closed by the front end subnet N_s, called artifacts removal stage. In this stage, a DCNN maps the digitized old halftone print Y to its synthetic counterpart $\hat{Y}$, aiming to remove common blemishes found in old halftone prints mentioned above. Due to the lack of real paired data, subnet N_s requires unsupervised learning. However, since the mapping between Y and $\hat{Y}$ is quite simple, unsupervised learning can solve this problem to the needed precision. Therefore, by dividing the original problem into two stages and tackling them individually, the proposed technique can effectively restore high quality images from old halftone prints.

The remainder of this paper is organized as follows. Section 2 overviews related object counting techniques. Section 3 discusses the proposed two-stage strategy. Section 4 presents evaluates the our technique with different experiments and Section 5 concludes.

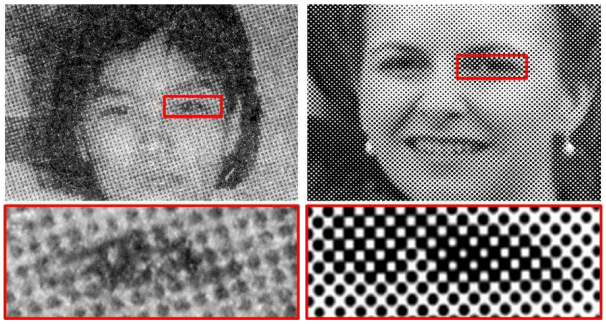
2. Related Work

Fundamentally, halftoning is a resampling process that transforms the grayscales of an analog signal to the duty cycles of a binary signal. Although the process removes most of the high frequency components of the original image, the low frequency components are well preserved in general. Thus, the most common conventional strategy of inverse halftoning is to use lowpass filters, such as edge-preserving filtering [29, 26, 21]. Xiong *et al*. proposed a wavelet domain filtering technique for reducing digital halftone pattern generated using error diffusion [43]. This technique employs cross-scale correlations to improve the preservation of edges during the removal of halftone artifacts. In [32], Siddiqui *et al*. adopted a modified locally adaptive SUSAN filter to suppress halftone patterns. They also developed an efficient local gradient based method called HFD, which exhibits good performances in smooth area but is ineffective around edges [33]. In [36], Sun *et al*. employed BM3D [6] to remove artifacts in halftone prints. This method also uses adaptive Guassian filtering to further smooth the image and bilateral filtering to improve the sharpness of edges. Ciobanu *et al*. proposed an artifact removal technique to clean up imperfections in scanned halftone prints as a preprocessing step for halftone restoration [5].

Statistical learning is also widely adopted in many inverse halftoning techniques, such as look-up table (LUT) [4], least-mean-square (LMS) filtering [3], maximum a posteriori (MAP) [35] and sparse representation [34]. Mese *et al*. proposed a recursive template selection algorithm for vector table generation and pixel value estimation using a constructed LUT [28]. Stevenson *et al*. proposed a MAP-based inverse halftoning technique that utilizes image distribution prior defined in Markov random field [35]. In [49], Zhang *et al*. presented a structure based multi-dictionary learning method, which employs feature dictionaries learned from image patch groups. In [16], Huang *et al*. presented a radial-basis function neural network (RBFNN) for inverse halftoning images. Pelcastre *et al*. proposed a method that use a multilayer neural network to improve the image quality of lowpass filtered halftone image [30].

Recently, DCNN has shown its great potential in image restorations and has achieved the state of the art for various problems including inverse halftoning. For example, for error diffusion based digital halftone pattern, ResNet and U-Net based inverse halftoning technique have been studied in a few researches [41, 14, 42]. Some of these proposed techniques adopt more complex perceptual constraints [18] to improve the visual quality of their outputs. For conventional dot-shaped halftone pattern, Kim *et al*. proposed a DCNN technique that can incorporate more context information by using a segmentation subnet and an edge detection subnet [19]. Overall, the performance of these methods is highly dependent on the quality of the training halftone samples. Using only synthetic halftone images for training, these methods cannot properly handle the various degradations in real halftone images scanned from old publications, resulting inaccurate and noisy outputs.

Besides the purposefully designed inverse halftoning DCNNs, some general purpose DCNN can also be retrained for the inverse halftoning task. For example, style transfer networks, which map an input image to a different style without requiring strictly paired training data [9], can transform a halftone image to a regular clean image given sufficient samples of such images. GAN [10], which is widely adopted to bridge the gap between different distributions in the learning of end-to-end mapping, is also helpful in alleviating the problem of lack of paired training data for inverse



Vintage halftone print Synthetic halftone image

Figure 2: It is difficult to simulate real halftone prints precisely, as they are often plagued by various types of blemishes such as ink smear, paper wear, etc.

halftoning. Many recent image restoration techniques have incorporated GAN in their designs [46, 8, 40]. Another interesting technique that can potentially solve the inverse halftoning problem is CycleGAN [50, 44]. CycleGAN employs two image-to-image translation networks of opposite directions to generate and validate paired training data from unpaired data. In theory, CycleGAN is capable of solving image restoration problems without given an explicit model of the degradation.

3. Two-Stage Restoration

As we discussed previously, the main challenge of using deep learning techniques for inverse halftoning is the lack of paired data. As low quality halftone prints in old publications are often the only surviving copies of the photographs, it is impossible to build a sufficiently large training set of real image pairs. One common solution to the problem is to use synthetic data for training [14, 19]. The main drawback of this approach is that synthetic data trained DCNNs tend to only work for synthetic input. Due to the vastly different types of blemishes plaguing real old halftone prints as exemplified in Figure 2, it is very difficult to simulate the complex compound effects of multiple degradation causes to the desired precision. As a result, the performance of synthetic data trained techniques often deteriorate significantly when dealing with real halftone images. Another solution is to use unsupervised learning [17, 50]. While this approach does not need strictly paired training data, a trained network might not be able to accurately learn the semantic connection between the degraded images and clean images, resulting various types of artifacts in the output.

Our approach is a mixture of the two aforementioned solutions. Instead of using a direct end-to-end mapping network N from real halftone print Y to its continuous tone counterpart X, we adopt a strategy of two-stage deep learning. The key idea of this strategy is to add a new intermediate objective, synthetic halftone image $\hat{Y}$, splitting the original mapping network $N\colon Y \rightarrow X$ into two cascaded subnets $N_s\colon Y \rightarrow \hat{Y}$ and $N_i\colon \hat{Y} \rightarrow X$. The first stage subnet N_s is the artifacts removal subnet, which maps the input real halftone image Y to its synthetic counterpart $\hat{Y}$. Despite the lack of strictly paired data, N_s is easy to train using unsupervised learning techniques, as removing artifacts from the real halftone prints into synthetic clean images is quite straightforward. The second stage subnet N_i is the inverse halftoning subnet, which maps the input synthetic halftone image $\hat{Y}$ to its continuous tone counterpart X. Since sufficient paired images can be easily generated using the halftone synthesizer M, N_i is easy to train as well. Therefore, by dividing the original direct end-to-end mapping, we break the restoration task into two much more tractable sub-problems.

3.1. Architecture

The goal of the first stage halftone artifacts removal subnet N_s is to remove common blemishes found in old halftone prints, making the resulting image statistically close to synthetic halftone prints. The subnet should only fix the defects of the halftone print and reconstruct the ink dots, i.e., the shape, size and grayscale of the dots, leaving the content of the image unchanged. To achieve such an objective, subnet N_s does not need to have a large number of parameters, but a sufficiently large receptive field is necessary for preserving halftone patterns. As shown in the left part of Figure 3, the architecture of the artifacts removal subnet is based on autoencoder [13, 7]. The encoder section of the network removes the details of the original halftone pattern and transforms the image content into compact feature representations. The details are then reconstructed in the decoder section according to the image content using the ideal halftone patterns found in synthetic halftone images. Both the encoder and decoder sections are consist of 8 convolution layers with stride 2 and kernel size 4. In order to decrease the unnecessary information extraction, skip connections [12, 27, 17] are employed in corresponding layers between the encoder and decoder.

The goal of the second stage inverse halftoning subnet N_i is to remove synthetic halftone pattern and reconstruct a high-quality image. The basic architecture of this subnet is shown in the right part of Figure 3. During the halftoning process, the low frequency components of the original image are mostly preserved, while the high frequency components are heavily distorted. The semantic information of the original high-quality image can only be revealed by examining the big picture. Thus, we add skip connections at the beginning of the subnet in order to provide a path for gradient propagation in optimization, reducing the difficulty of training. We also use two strided convolutional layers for expanding the receptive field. Inspired by the dense block

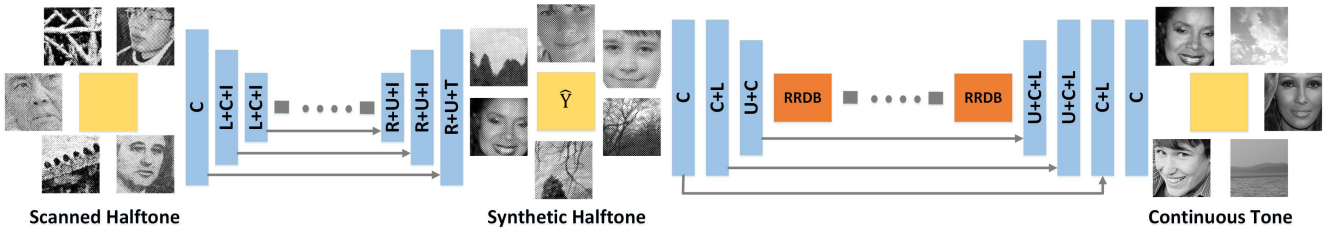


Figure 3: The network architecture of the proposed technique. Symbols C, U, I, L, R, T denote convolution layer, transposed convolution layer, instance normalization, LeakyReLU ($\alpha = 0.2$), ReLU, and Tanh respectively

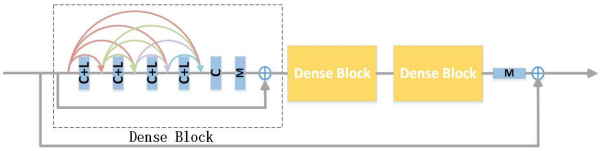


Figure 4: RRDB block. Symbols C, L, M denote convolution layer, LeakyReLU ($\alpha = 0.2$), residual scaling, respectively

[15, 48, 40], we increase network capacity using dense connection. High frequency information prediction is based on the abstract context feature space with 16 stacked residual in residual dense blocks (RRDB) as shown in Figure 4. Residual scaling [24, 37] is also employed for preventing instability.

3.2. Objective Functions

In the training of the proposed technique, the two subnets N_s and N_i are first trained separately using different objective functions. Then we combine the two subnets together, forming the complete network as shown in Figure 3, and train them jointly.

Objective function for halftone artifacts removal. The mapping function N_s from Y to $\hat{Y}$ is trained using the GAN technique, where the adversarial loss is defined as:

$$\mathcal{L}_{N_s} = \mathbb{E}_Y \left[-\log(D(N_s(Y))) \right] \tag{1}$$

Discriminative network D is trained jointly with N_s [10]. To improve the performance of the network and stabilize the training, we add an inverse mapping network N'_s from $\hat{Y}$ to Y sharing the same architecture as N_s to form a cycle as in CycleGAN [50]. In addition to the adversarial loss of the inverse mapping network, cycle-consistency loss is also employed, which is defined as,

$$\begin{aligned} \mathcal{L}_{\text{cycle}} = \mathbb{E}_Y[\|N'_s(N_s(Y)) - Y\|_1] \\ + \mathbb{E}_{\hat{Y}}[\|N_s(N'_s(\hat{Y})) - \hat{Y}\|_1] \end{aligned} \tag{2}$$

The overall objective function for the halftone artifacts removal subnet is as follows,

$$\mathcal{L}_{\text{stage1}} = \mathcal{L}_{N_s} + \mathcal{L}_{N'_s} + \mathcal{L}_{\text{cycle}} \tag{3}$$

Objective function for inverse halftoning. In this stage, large quantities of paired synthetic data $\{\hat{Y}, X\}$ are available. Thus, the training can be carried out using straightforward end-to-end supervised learning. As in many other DCNN-based image restoration techniques, we use L_1 norm as the objective function as follows,

$$\mathcal{L}_{\text{stage2}} = \mathbb{E}_{\hat{Y},X} \|N_i(\hat{Y}) - X\|_1. \tag{4}$$

Objective function for joint training. To further improve the performance of the technique, we jointly fine-tune subnets N_s and N_i after completing their individual training. Since there is no real paired training data $\{Y, X\}$ of real halftone prints and their high-quality counterparts, simple end-to-end supervised learning is infeasible for this task. To solve this problem. we borrow the idea of deep learning with surrogate ground truth [25]. For each input real halftone prints Y, we generate a surrogate ground truth image $S(Y)$ using a filtering-based inverse halftoning operator called Sattva [1]. While Sattva preserves low frequency components of the input image very well, it often fails to reconstruct the high frequency components to desired precision for real halftone prints. Thus, in the surrogate loss, only the low frequency components of the results of the joint network N are compared against the surrogate ground truth, as follows,

$$\mathcal{L}_s = \mathbb{E}_Y \|\mathcal{F}(N(Y)) - \mathcal{F}(S(Y))\|_1, \tag{5}$$

where $\mathcal{F}$ is a function that removes high frequency information of the input image. In all of the presented experiments, we use a pre-trained 19-layer VGG network as $\mathcal{F}$. To prevent over-smoothing and to generate photorealistic results, adversarial loss $\mathcal{L}_{\text{adv}}$ is also added in the joint training. By combining the loss used in the previous individual training, we arrive at the final formulation of the joint training objective function,

$$\mathcal{L}_{\text{joint}} = \lambda_1 \mathcal{L}_{\text{stage1}} + \lambda_2 \mathcal{L}_s + \lambda_3 \mathcal{L}_{\text{adv}} \tag{6}$$

where λ_1, λ_2 and λ_3 are the weights balancing the different objective terms.

4. Experiments

To evaluate the performance of the proposed two-stage technique, we implement the DCNN-based algorithm using PyTorch and compare it against several other popular techniques for inverse halftoning. In this section, we discuss the details of the experiments.

4.1. Datasets and Training Details

The ground truth images for the training of the inverse halftoning subnet N_i are collected from several popular image datasets, including DIV2K [2], OutdoorScene [39], Helen [23], IBUG [31], UTKFace [47]. From the total 12436 high quality images collected, 8700 are randomly selected for training. The corresponding synthetic halftone images are generated from these images by the aforementioned halftone image synthesizer with random halftone scales and angles. Using real vintage halftone prints as references, the distances of adjacent halftone dots are set between 3 and 9; and angles of the patterns are set between 15° and 75° in the our experiments. For evaluating the halftone artifacts removal subnet and the joint network, 379 vintage photo prints are scanned from old newspapers and books, where 195 images are used for training.

The two subnets N_s, N_i are first trained separately for 120 and 500 epochs, respectively. In both cases, the learning rate is set to 0.0002 for the first half of the training process and then decreased to 0 linearly. Kaiming initialization is adopted for all the networks [11]. Adam optimizer is used in optimization with $\beta_1 = 0.5$ [20]. After the training of the two subnets, the whole network is trained jointly for another 120 epochs at learning rate 0.00005. The discriminator of N_s shares a similar architecture as the discriminator of DCGAN. Their main difference is that the proposed method uses instance normalization instead of batch normalization. The discriminator of N_i employs a VGG-style architecture but with max pooling replaced by strided convolution. For the joint training objective function as in Eq. (6), the weights $\lambda_1, \lambda_2, \lambda_3$ are empirically set to 1, 0.5 and 1, respectively. All the experiments are carried out on a computer with a NVIDIA GTX Titan X GPU.

4.2. Baseline Algorithms

To evaluate the performance of the proposed technique, the following algorithms are also tested for comparison.

Sattva [1]. This is a low-pass filtering based method implemented as a plugin for Photoshop. It automatically detects the density of halftone, and apply low-pass filters accordingly to prevent over-smoothing and moiré patterns. Sattva is one of the best conventional inverse halftoning technique in terms of visual quality.

ESRGAN [40]. This is originally a state-of-the-art learning based single image super-resolution algorithm. To repurpose ESRGAN as an inverse halftoning technique, we first downsample the given halftone image by a factor of 4 using a Guassian kernel, and then feed the result to ESRGAN to restore the image to the original scale. In general, this method performs much better than direct end-to-end mapping networks, even with the help of GAN. The reason is that, since most high-frequency components of the original image are severely distorted during the halftoning process, the fine details of a real halftone image, including paper fiber structure, ink smear, etc., offer little or no useful information and can often become a source of interference. Downsampling greatly reduces this problem, making the ESRGAN method more effective for inverse halftoning. To further improve the baseline performance, we retrain the official implementation of ESRGAN using our dataset.

CycleGAN [50]. This is an unpaired image-to-image translation DCNN. Since CycleGAN learns a mapping from source domain to target domain without the necessity of paired training samples, it could solve the problem of the lack of ground truth image for vintage halftone prints. We trained CycleGAN directly using unpaired real halftone images and continuous tone images from our dataset.

Unet [14] and **DescreenNet** [19]. Both are learning based inverse halftoning method that produce excellent results for synthetic halftone images. The implementation of Unet [14] is from the original authors. We fine-tuned their pre-trained model using our dataset. For DescreenNet [19], since its implementation is not made available publicly, we only compare the results published by the authors in the our evaluation.

	ESRGAN	Unet	Proposed
PSNR	29.94	28.60	**30.10**
SSIM	0.8904	0.8561	**0.8951**

Table 1: Average PSNR and SSIM performances of the tested techniques on synthetic data.

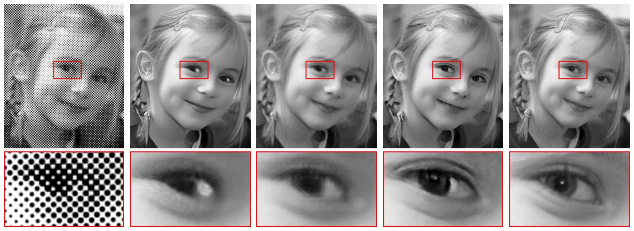


Figure 5: Sample results of algorithms on synthetic image.

4.3. Comparison of Algorithms

Shown in Table 1 are the quantitative evaluation results of the tested algorithms on synthetic dataset. For fair com-

	Hafltone	Sattva	ESRGAN
Accuracy	0.6012	0.7484	0.7086
	CycleGAN	Unet	Proposed
Accuracy	0.7465	0.7331	**0.7576**

Table 2: Accuracy of the MTCNN face detection algorithm on images restored by inverse halftoning techniques.

parison, adversarial optimization is disabled for ESRGAN and the proposed method. As shown in the table, the proposed method achieved the best result, but the differences are insignificant. Some samples of the synthetic test results are presented in Figure 5. For synthetic halftone images, the proposed method exhibits more details than the competition. The results from ESRGAN and Unet look blurry in comparison.

When it comes to real scanned halftone prints, the performance of the proposed algorithm is far superior than the other tested algorithms. Figure 6 shows the results on real halftone prints provided by the authors of DescreenNet [19]. As shown in the figure, the proposed method produces sharper and more natural looking details than Sattva and the other DCNN-based approaches. The proposed algorithm also works well on our scanned halftone print set, which includes portraits and outdoor scenery, as shown in Figure 7. In comparison, our algorithm produces much cleaner and refined results. It also generates more realistic looking details than the other DCNN approaches. The results of the compared techniques are often plagued by high frequency noise and unpleasant artifacts. More experimental results and comparison with other methods are available in the supplementary material.

Since the originals of most real vintage halftone prints have long been lost in the past, it is difficult to evaluate the performance of inverse halftoning techniques on real halftone prints objectively without the ground truth. Therefore, instead of using direct quantitative measurement, we evaluate the performance of face detection on images restored by the inverse halftoning methods as a circumstantial evidence of the restoration quality. Shown in Table 2 is the average face recognition accuracy of pre-trained MTCNN [45] on the input halftone prints and the corresponding results by the tested inverse halftoning methods. MTCNN achieves the best performance when its input image is the result of the proposed method. This shows that the proposed method produces the least amount of artifacts on halftone portraits in comparison with the other tested techniques.

4.4. Ablation Study of the Proposed Method

In this subsection, we test various ablations of our full architecture to evaluate the effects of each components of the proposed algorithm.

Artifacts removal subnet. Figure 8b shows the results of the proposed algorithm without the artifacts removal subnet. While the results look realistic in coarse scale, lots of the imperfections, such as paper fiber patterns and blemish, are carried over from the input halftone prints, making the results look noisy in fine scale. The results of ESRGAN and Unet also exhibit this problem, as shown in Figure 7.

Inverse halftoning subnet. Figure 8c presents the results of the proposed algorithm without the inverse halftoning subnet. As demonstrated in the figure, the results are rich in detail, but they look unrealistic overall and contain severe artifacts in smooth regions, similar to the results of CycleGAN as shown in Figure 7. Although it is still possible to train a DCNN using unpaired data in the absences of the inverse halftoning subnet, it is too difficult to make it learn the complex and compounded degradation of vintage halftone prints precisely. The intermediate synthetic halftone image provides an important guide, greatly reducing the difficulty of the problem.

Joint training. Figure 8d shows the results of the proposed algorithm without the final joint training. While most paper fiber and halftone patterns are successfully removed as shown, there are still some artifacts in smooth regions. In comparison, the results of the jointly optimized version show no sign of these problems, as demonstrated in Figure 8e.

4.5. Improving Existing Methods

We also investigate the possibility of improving existing inverse halftoning methods using our artifacts removal subnet N_s. As shown in Figure 9b, using Unet directly on real scanned halftone prints often leaves objectionable noise in the resulting image. By first pre-processing the input real halftone prints with the artifacts removal subnet N_s, the visual quality of the output image by Unet is greatly improved, as demonstrated in Figure 9c. Many similar inverse halftoning methods can benefit from a pre-trained subnet N_s when dealing with real vintage halftone prints.

4.6. Generating Realistic Halftone Images

Interestingly, with the inverse artifacts removal subnet N_s' used in the training of N_s, we can also generate photorealistic halftone image from clean synthetic halftone image, as exemplified in Figure 10. Thus, subnet N_s' can be utilized potentially as a more realistic halftone image synthesizer for future research of halftone image restoration.

4.7. Model Size and Computation Time

Summarized in Table 3 are the model sizes and computation times for processing an image of size of 512×512 with existing methods. The time cost of DescreenNet [19] reported by the original authors is also listed as a reference.

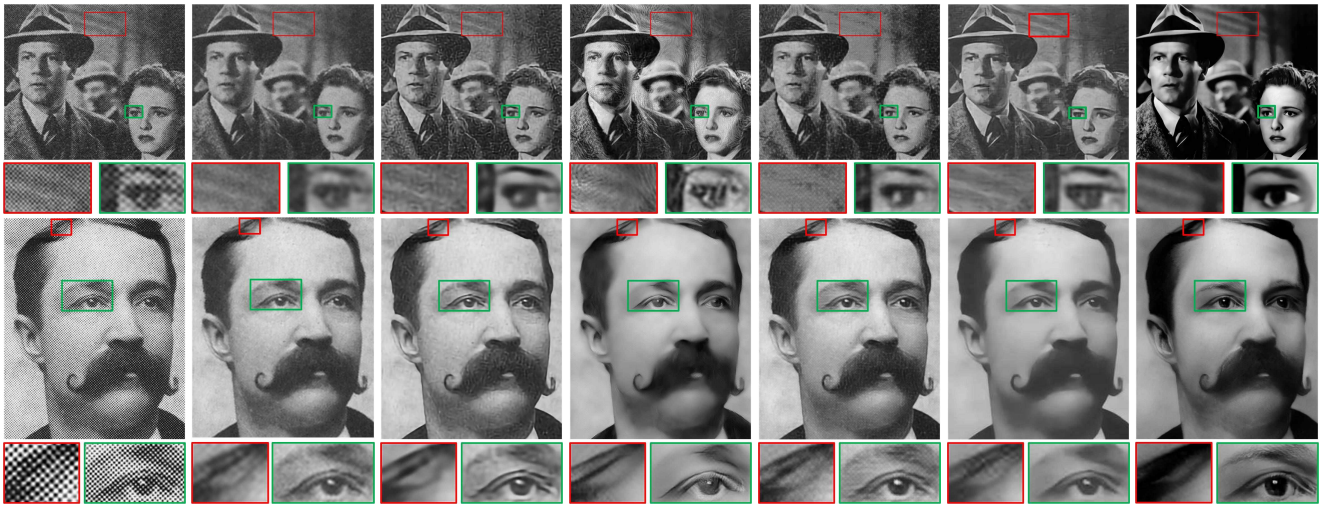


Figure 6: Sample results of algorithms on real halftone prints provided by [19].

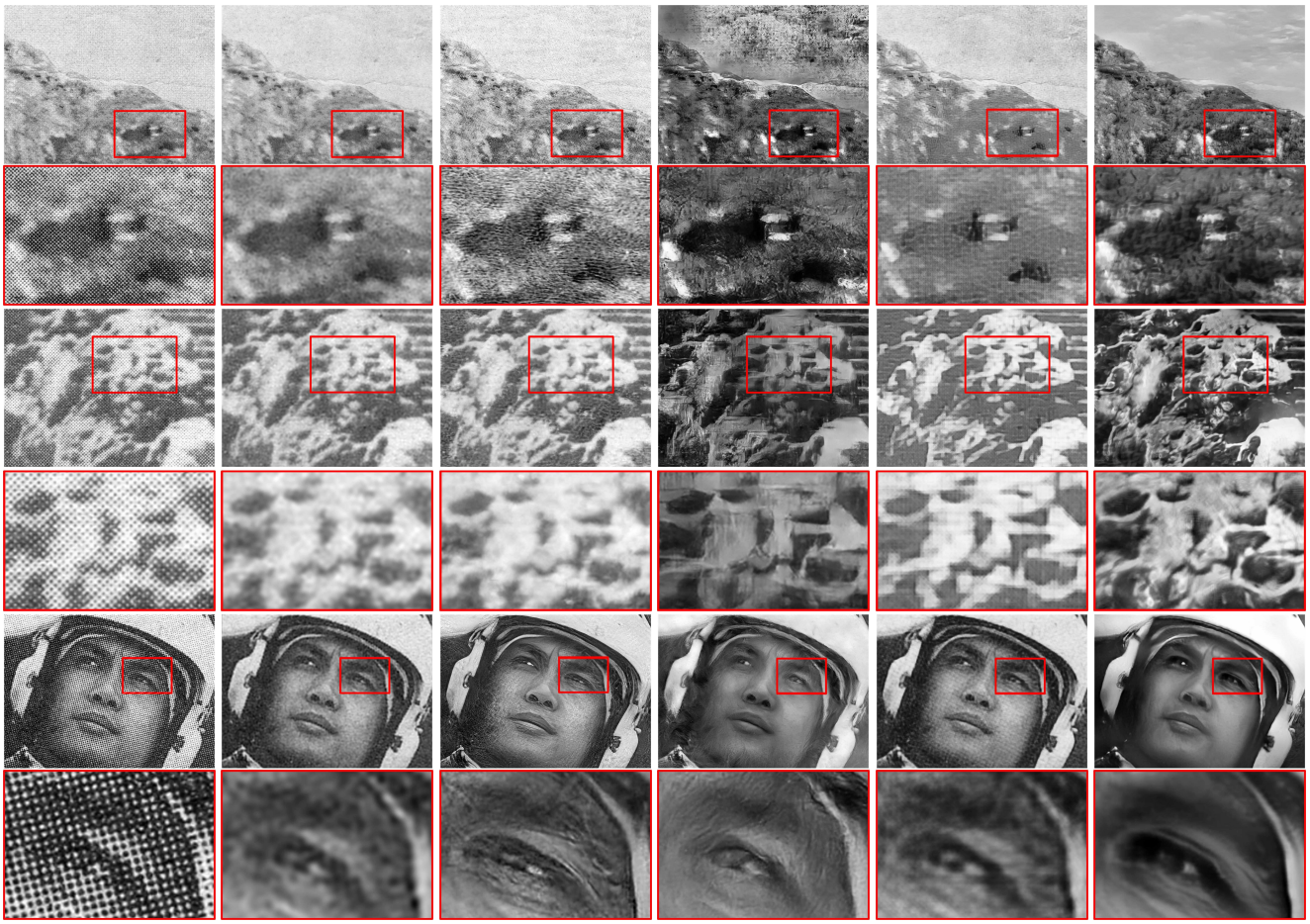
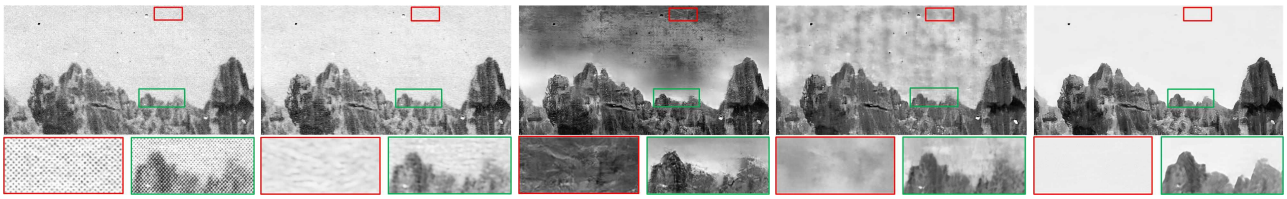
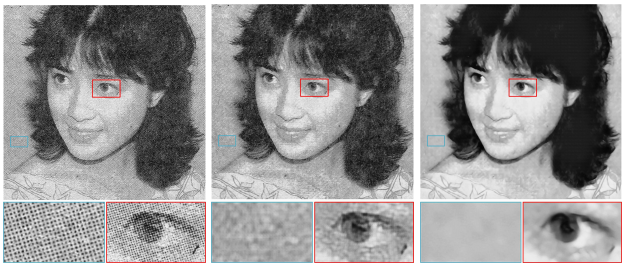


Figure 7: Sample results of algorithms on real halftone prints scanned from old publications.



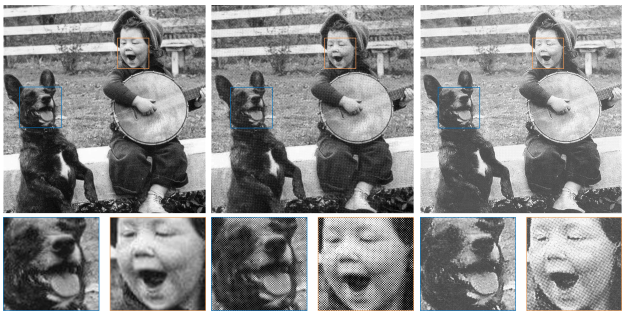
(a) Input (b) No artifacts removal (c) No inverse halftoning (d) No joint training (e) Full version

Figure 8: Ablation study of the proposed technique. The network performs the best when the artifacts removal and inverse halftoning subnets are both employed.



(a) Halftone print (b) Unet (c) N_s+Unet

Figure 9: By pre-processing the input halftone print with the proposed artifacts removal subnet N_s, the output quality of Unet is greatly improved.



(a) Input (b) Synthetic (c) Realistic

Figure 10: With the inverse artifacts removal subnet N_s', we can generate realistic halftone images.

4.8. Limitation

The proposed method might not be able to fully recover the original photos to the desired quality if the input halftone prints are severely damaged. As exemplified in Figure 11, the proposed method often fails to restore faded or poorly illuminated parts in a halftone print. It can also mistake ink spots as true signal and leave them in restored images. A possible solution to the problem is to add a preprocessing stage to detect ink spots or other types of imperfections. Then, we can use other restoration techniques, such as inpainting, to recover these regions separately.

Table 3: Comparison of model sizes and computation times for processing an input image of size 512×512 with existing methods.

Method	Time (s)	Parameters ($\times 10^6$)
Unet[14]	0.003	13.603
DNIH[42]	0.003	3.478
ESRGAN[40]	0.054	16.696
CycleGAN[50]	0.005	11.366
DescreenNet[19]	0.070	N/A
Proposed	0.042	71.283

Figure 11: Sample of failure cases. Some severse blemishes can disrupt the restoration.

5. Conclusion

In this research, we propose a novel strategy that divides the task of real halftone prints restoration into two stages: the removal of printing artifacts and the inverse of halftoning. The advantage of our technique is that only the simple first stage requires unsupervised training, while the more complex second stage of inverse halftoning can be easily trained with synthetic data. Extensive experimental results demonstrate the superiority of the proposed algorithm over the state of the art for vintage halftone prints.

References

[1] Sattva descreen. http://www.descreen.net/, 2019.

[2] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

[3] Li-Ming Chen and Hsueh-Ming Hang. An adaptive inverse halftoning algorithm. *IEEE transactions on image processing*, 6(8):1202–1209, 1997.

[4] Kuo-Liang Chung and Shih-Tung Wu. Inverse halftoning algorithm using edge-based lookup table approach. *IEEE Transactions on Image Processing*, 14(10):1583–1589, 2005.

[5] Adrian Ciobanu, Tudor Barbu, and Mihaela Luca. Image restoration for halftone pattern printed pictures in old books. In *International Conference of Electronics, Computers and Artificial Intelligence*, page 91. IEEE, 2018.

[6] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

[7] Gabriel Eilertsen, Joel Kronander, Gyorgy Denes, Rafał K Mantiuk, and Jonas Unger. Hdr image reconstruction from a single exposure using deep cnns. *ACM Transactions on Graphics (TOG)*, 36(6):178, 2017.

[8] Deniz Engin, Anıl Genç, and Hazım Kemal Ekenel. Cycle-dehaze: Enhanced cyclegan for single image dehazing. *arXiv preprint arXiv:1805.05308*, 2018.

[9] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2414–2423, 2016.

[10] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer, 2016.

[13] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

[14] Xianxu Hou and Guoping Qiu. Image companding and inverse halftoning using deep convolutional neural networks. *arXiv preprint arXiv:1707.00116*, 2017.

[15] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *CVPR*, volume 1, page 3, 2017.

[16] Win-Bin Huang, Alvin WY Su, and Yau-Hwang Kuo. Neural network based method for image halftoning and inverse halftoning. *Expert Systems with Applications*, 34(4):2491–2501, 2008.

[17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, 2017.

[18] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer, 2016.

[19] Tae-Hoon Kim and Sang Il Park. Deep context-aware descreening and rescreening of halftone images. *ACM Transactions on Graphics (TOG)*, 37(4):48, 2018.

[20] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[21] Thomas D Kite, Niranjan Damera-Venkata, Brian L Evans, and Alan C Bovik. A fast, high-quality inverse halftoning algorithm for error diffused halftones. *IEEE Transactions on Image Processing*, 9(9):1583–1592, 2000.

[22] Dorothy M Kosinski and Dorothy M Kosinski. *The artist and the camera: Degas to Picasso*. Dallas Museum of Art Dallas, TX, 1999.

[23] Vuong Le, Jonathan Brandt, Zhe Lin, Lubomir Bourdev, and Thomas S Huang. Interactive facial feature localization. In *European conference on computer vision*, pages 679–692. Springer, 2012.

[24] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *The IEEE conference on computer vision and pattern recognition (CVPR) workshops*, volume 1, page 4, 2017.

[25] Bolin Liu, Xiao Shu, and Xiaolin Wu. Deep learning with inaccurate training data for image restoration. *arXiv preprint arXiv:1811.07268*, 2018.

[26] Jiebo Luo, Ricardo De Queiroz, and Zhigang Fan. A robust technique for image descreening based on the wavelet transform. *IEEE Transactions on Signal Processing*, 46(4):1179–1184, 1998.

[27] Xiaojiao Mao, Chunhua Shen, and Yu-Bin Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In *Advances in neural information processing systems*, pages 2802–2810, 2016.

[28] Murat Mese and Palghat P Vaidyanathan. Look-up table (lut) method for inverse halftoning. *IEEE Transactions on Image Processing*, 10(10):1566–1578, 2001.

[29] Christopher M Miceli and Kevin J Parker. Inverse halftoning. *Journal of Electronic Imaging*, 1(2):143–152, 1992.

[30] Fernando Pelcastre-Jimenez, Mariko Nakano-Miyatake, Karina Toscano-Medina, Gabriel Sanchez-Perez, and Hector Perez-Meana. An inverse halftoning algorithm based on neural networks and up (x) atomic function. In *Telecommunications and Signal Processing (TSP), 2015 38th International Conference on*, pages 523–527. IEEE, 2015.

[31] Christos Sagonas, Epameinondas Antonakos, Georgios Tzimiropoulos, Stefanos Zafeiriou, and Maja Pantic. 300 faces

in-the-wild challenge: Database and results. *Image and vision computing*, 47:3–18, 2016.

[32] Hasib Siddiqui and Charles A Bouman. Training-based descreening. *IEEE transactions on image processing*, 16(3):789–802, 2007.

[33] Hasib Siddiqui, Mireille Boutin, and Charles A Bouman. Hardware-friendly descreening. *IEEE Transactions on Image Processing*, 19(3):746–757, 2010.

[34] CH Son and HM Park. Sparsity-based inverse halftoning. *Electronics letters*, 48(14):832–834, 2012.

[35] Robert L Stevenson. Inverse halftoning via map estimation. *IEEE Transactions on Image Processing*, 6(4):574–583, 1997.

[36] Bin Sun, Shutao Li, and Jun Sun. Scanned image descreening with image redundancy and adaptive filtering. *IEEE Transactions on Image Processing*, 23(8):3698–3710, 2014.

[37] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI*, volume 4, page 12, 2017.

[38] Michael Twyman and Ruari MacLean. *Printing 1770-1970: an illustrated history of its development and uses in England*. British library London, 1998.

[39] Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[40] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *The European Conference on Computer Vision Workshops (ECCVW)*, September 2018.

[41] Menghan Xia and Tien-Tsin Wong. Deep inverse halftoning via progressively residual learning. In *Asian Conference on Computer Vision*, page 59, 2018.

[42] Yi Xiao, Chao Pan, Xianyi Zhu, Hai Jiang, and Yan Zheng. Deep neural inverse halftoning. In *2017 International Conference on Virtual Reality and Visualization (ICVRV)*, pages 213–218. IEEE, 2017.

[43] Zixiang Xiong, Michael T Orchard, and Kannan Ramchandran. Inverse halftoning using wavelets. *IEEE transactions on image processing*, 8(10):1479–1483, 1999.

[44] Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong. Dualgan: Unsupervised dual learning for image-to-image translation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2849–2857, 2017.

[45] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, Oct 2016.

[46] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Learning a single convolutional super-resolution network for multiple degradations. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 6, 2018.

[47] Song Yang Zhang, Zhifei and Hairong Qi. Age progression/regression by conditional adversarial autoencoder. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2017.

[48] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[49] Yan Zhang, Erhu Zhang, Wanjun Chen, Yajun Chen, and Jinghong Duan. Sparsity-based inverse halftoning via semi-coupled multi-dictionary learning and structural clustering. *Engineering Applications of Artificial Intelligence*, 72:43–53, 2018.

[50] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networkss. In *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.