Digital Multitone Image Reconstruction using Deep Generative Adversarial Nets

Jing-Ming Guo* and Sankarasrinivasan S†
*†National Taiwan University of Science and Technology
* E-mail: jmguo@seed.net.tw Tel: +886-27303241
†E-mail: D10507805@mail.ntust.edu.tw Tel: +886-978374981

Abstract— Digital multitoning is a latest version of the halftone technology which is widely used in printing applications. The method transforms the original gray scale image into its appropriate printable version using more tones than conventional halftones. To be specific, the gray scale or color image is converted to its 2 or 3-bit version. The multitone image results in enhanced homogeneity, smooth texture and less halftone artifacts. On the other hand, many applications require reconstruction of printed images for digital processing and in most cases source image may not be available. In this paper, a deep learning model is proposed based on the conditional general adversarial nets (cGAN's) to perform effective reconstruction of multitone images. For experimentation, a digital multitone dataset is developed comprising of 10,000 images. The proposed model is optimized to perform with minimal layers and reduced parameters. The discriminator is constructed as Patch-GAN architechture and is implemented over the patch size of 64x64. From the overall experimental results, it has been validated that the proposed model achieves consistent and best reconstruction quality.

Keywords: Digital multitoning, GAN, Image reconstruction, Inverse halftoning, Printing.

I. INTRODUCTION

Digital halftoning is a classical transformation technique to convert the gray scale image into its equivalent bitmap (binary) version [1]. When the bitmap image is printed and viewed from the distance it is perceived as a continuous-tone image. Digital multitoning is a newly developed halftoning technique, which involves converting the gray-scale image (8-bit) version into its lower-bit version [2]. Typically, in halftoning the gray-scale image is converted into the 2-bit version using standard black cartridge, whereas multitoning use additional tonal cartridges. The multitone image has superior image quality than the conventional halftone algorithms.

In earlier days, the conventional halftone algorithms utilize ordered dithering techniques [3] using the standard clustered and dispersed dot dither arrays. Though it is a straight forward and very simple approach, the output binary image suffers from visible artifacts and non-homogeneous pattern. Furthermore, the error and dot diffusion [4] are introduced which involves distributing the error in the neighborhood pixels to improvise the rendered image quality. Finally, direct binary search (DBS) [6] technique gained prominence as it can deliver superior quality images and its operated based on Human Visual System (HVS) filter. At the same instant, as DBS works on iterative strategy it takes high processing time and practically difficult to implement in hardware. The major break though is achieved, with the introduction of blue and green noise based dither array, which can provide superior quality images in comparison with DBS techniques [6]. With this extensive research in optimization of halftone patterns, the quality of the two tone print output is almost saturated and very difficult for further improvisations. In subsequent developments, Digital multitone strategy is introduced in order to overcome the limitations of conventional halftones and it can offer superior quality print output. The digital multitoning is broadly classified based on the spectral properties such as blue and green noise to support for various printers. Blue noise pattern is preferable for inkjet printers which can deliver consistent rendering of various shades and stable dot patterns. Moreover, laser printers adopt green noise properties that suits its nature of unstable dot printing. In many application, the reconstruction of printed images is very important to digitally enhance and publish them.

Among the various deep learning models for reconstruction, the image-to-image (I2I) translation algorithm is considered and optimized with respect to the multitone reconstruction. As the I2I model relies on the generative adversarial network (GAN) to learn the loss function, the method is observed to have better image quality. The I2I model is based on the conditional GAN [7] which requires that the network need to be trained with the with an original image, and multitone image along with the noise as shown in Fig.1.

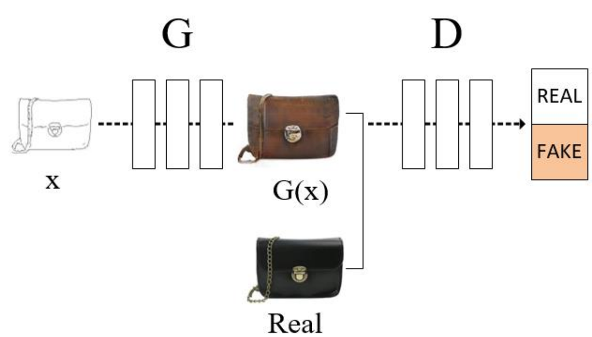


Fig. 1. Conditional Generative Adversarial Networks (cGAN's).

From the literature, it has been validated that the cGAN's are very effective in many image transformation problems such as

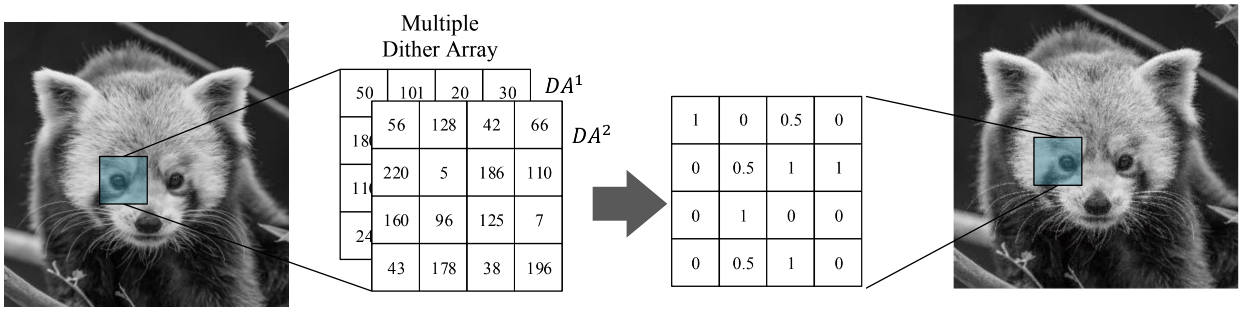


Fig. 2. Construction of Multitone Image

edge to image, map to satellite view, image to cartoon, text to image and so on.

As the approximate information is available in the image reconstruction problem, the present study also focusses on optimizing the architecture.

The main contribution of the paper are as follows:

- A digital multitone database is constructed for the clustered-dot types with different block sizes ranging from 8 to 64. The database comprises of 10,000 multitone images.
- A reduced conditional GAN architechture is proposed to expedite the training and to reduce the parameters significantly. The number of convolutional layers and filters are reduced without compensating the performances.
- The loss function is also modified with respect to the image reconstruction problem. The original cGAN's utilize L1 loss, and its weightage factor is reduced to improvise the structural information reconstruction.

II. Digital Multitoning

Multitoning comprise of sequence of block-wise thresholding using different level of dither array as illustrated in Figure 2. Each block output will be assigned with either 1 or 0, or its intermediate values. For example, 3 and 4 Tone is performed as follows,

For example, 3 Tone Multitoning is performed as follows,

$$Q_{B3} = \begin{cases} 1 & \text{if im [m, n]} > DA^1 \\ 0.5 & \text{if } DA^1 \geq \text{im [m,n]} > DA^2 \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The 4 Tone Multitoning is performed as follows,

$$Q_{B4} = \begin{cases} 1 & \text{if im [m, n]} > DA^1 \\ 0.66 & \text{if } DA^1 \geq \text{im [m,n]} > DA^2 \\ 0.33 & \text{if } DA^2 \geq \text{im [m,n]} > DA^3 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

DA^1, $DA^2, DA^3 \ldots \ldots, DA^{NT-1}$ be the clustered dot or Direct Multi Bit Search multitone dither arrays. NT- represents number of tones. Fig. 1 shows the multitone output generated using the elaborated dispersed and clustered dot screens.

III. Digital Multitone Image Reconstruction

Image to image translation [7] often deals with the image classification, reconstruction or regression problem that operates in a per-pixel level. As the input and output pixels are mapped in one to one basis, they are assumed to be conditionally independent over each other. Usually the network relies on two loss function such as L_1 and a structural loss element. The main role of the conditional GANs is to learn the structural loss L_{cGAN} during the training which enhances the realistic nature of the reconstructed image.

The conventional generator networks deal with mapping a random noise vector to an expected output image. On the other hand, Condition GAN learn the mapping between an observed image x and a random noise vector z to the output image. Relevantly, the objective function can be defined as

$$L_{cGAN}\,(G, D) = E_{x,y}[\log D(x, y)] + E_{x,z}[\log(1 - D(x, G(x, z)))] \quad (3)$$

During the optimization, the G tries to minimizes the function whereas the D tries to maximize the function. In addition to this, integration of L_1 or L_2 loss is also proven to be an effective strategy and the final objective function is provided as

$$G^* = \arg min_G\, max_G\, L_{cGAN}\,(G, D) + \lambda\mathcal{L}_{L1}(G) \quad (4)$$

The added output noise vector is of gaussian type.

Usually the L_1 or L_2 loss mainly suppress the high frequency elements, and hence the image output usually is blurred in nature. In general, the model rely on the L_1 loss to accurately capture the low frequency information and restricting the GAN architecture to only to the high frequency information that is related to the structural information in the image patch.

The discriminator operates as a Patch-GAN which penalizes the structural information at the scale of the patches, and the discriminator tries to determine whether the output image patch of size NxN is real or fake. In this approach, the image pixels are treated as a Markov random field elements and the assumption regarding the independency between the pixels is established based on the patch diameter.

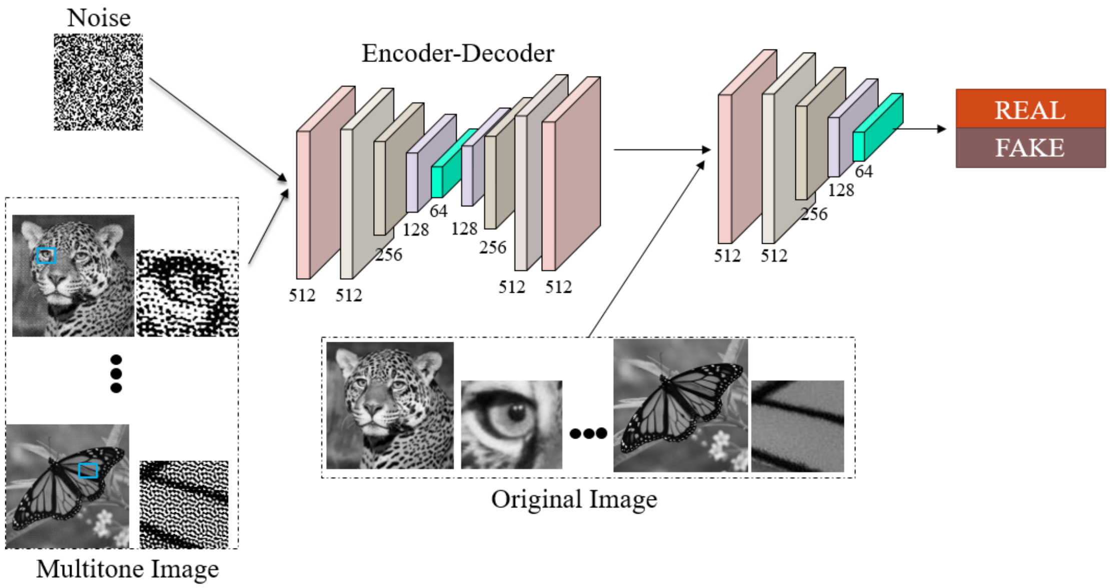


Fig. 3. Proposed C-GAN for Multitone Reconstruction

The generator model comprises of an encoder-decoder architecture which consist of Convolutional layers are used along with the batch normalization technique. The ReLu layers is used as the activation function layer. The filter size of the convolution layers is 4x4 with the stride rate of two.
The number of filters for each convolutional layer of the encoder and decoder is provided as

Encoder: CON64-CON128-CON256-CON512-CON512

Decoder: CON512-CON512-CON256-CON128-CON64

The number of filters for the discriminator architecture is

Discriminator: CON64-CON128-CON256-CON512-CON512

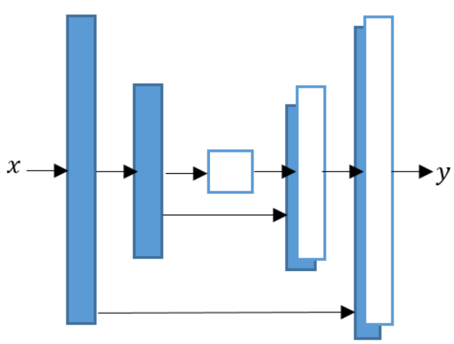


Fig. 4. Encoder-Decoder with skip connection

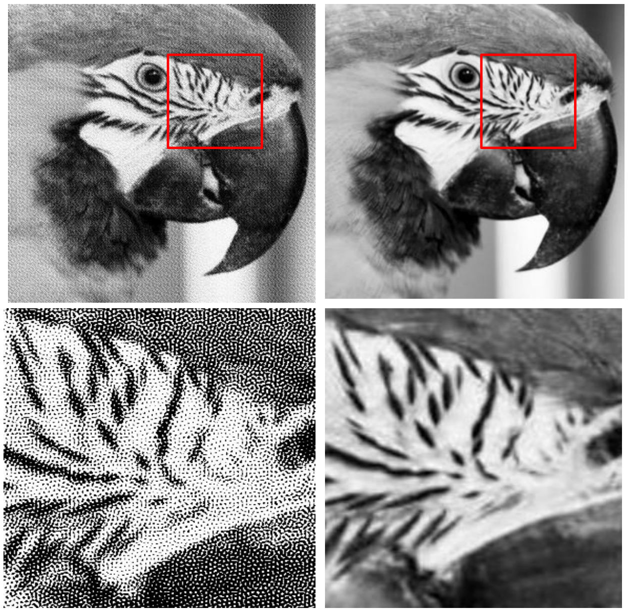
Figure 4, illustrates the architecture of the generator in which the encoder-decoder is provided with a skip connection. The initial weights are initialized based on the Gaussian distribution with mean 0 and standard deviation 0.01.

As the multitone images are well approximated with respect to the given image, the main challenge in restoration is to remove the inherent halftone and impulse noise associated with halftoning and to improvise the structural similarity. From our experiments, it has been inferred that the L1 loss is not contributing significant impact to the visual quality as low pass information is already available and hence the L1 loss term is assigned with less weightage (=10). And as image patches with high structural information are going to play a key role in loss function optimization, the GAN network is optimized considering only the effective patches. The effective patches are estimated using the variance and this also results in faster learning. The discriminator works as a PatchGAN in which it tries to predict the generated image in a patch wise manner. The patch size of the discriminator is set in accordance with the block size over which the multitone image are constructed. As the multitone is block-wise compression technique, the use of PatchGAN is very relevant and found to be very effective in obtaining a good reconstruction. For the generation of multitone image various block sizes has been applied such as 8, 16, 32 and 64. During the training, the algorithm is trained for multitone images of different sizes and accordingly the patch size diameter is also changed.

The number of images used for training and testing is 9000 and 1000 images. The algorithm is tested for the 100 epoch and the generator loss is observed to be minimal and stabilized. As the multitone with higher block sizes has more noises, the PatchGAN is set to 64x64 and it helps to reduce the noise significantly.

IV. RESULTS

For quality assessment, structural similarity index (SSIM) is used to evaluate the quality between the original multitone image and its reconstructed version. The reconstruction accuracy of the 1000 tested images is listed in Table I.



(a) Digital Multitone Image (b) Reconstructed Multitone Image

(c) Scanned Multitone Image (b) Reconstructed Multitone Image

Fig. 5. Reconstructed digital and scanned multitone image.

TABLE I. QUALITY OF RECONSTRUCTED MULTITONE IMAGE

Block Size	Multitone Image	C-GAN	Proposed C-GAN
8	0.6958	0.9536	**0.9612**
16	0.6433	0.9200	**0.9310**
32	0.6109	0.9005	**0.9061**
64	0.5842	0.8781	**0.8814**

Table I shows the results of improvements in the image quality of the reconstructed multitone image. It can be noted that the reconstruction rate is superior for the images with higher block sizes and it is due to the presence of heavy visible noises. The proposed method also performs better with respect to the original C-GAN architechture. Figure 5 illustrates the reconstructed image of digital and scanned multitone image. It can be clearly seen that the proposed methods had good reconstruction of structural information.

V. CONCLUSION

A deep learning based approach is proposed for the superior reconstruction of the multitone images. To begin with, the digital multitone database is constructed which comprise of around 10,000 images. The inherent noises associated with a multitone images such as halftone artifacts and impulsive noises is studied and Conditional General Adversarial Networks (cGAN's) are exploited to solve this problem. As the image to image translation architecture is formulated for the edge to image and many complex problems, the present work considerably reduced the number of filter and layers with respect to existing work. The loss function of the existing architecture is also fine-tuned to perform for multitone images, and the loss functions that corresponds to retaining low pass information is neglected. The PatchGAN based discriminator model is set to perform for block size of 64x64, and the decision over real or fake image is decided for this specific image size. From the results, it is validated that the proposed strategy obtained very good image quality and the visible inherent halftone based noises are eliminated significantly. In future work, the architecture is extended to perform for reconstruction of other compressed image versions.

REFERENCE

[1] D. L. Lau, and R. A. Gonzalo., Modern digital halftoning. CRC Press, 2008.

[2] Y. F. Liu et. al., "Clustered-dot screen design for digital multitoning," *IEEE Transactions on Image Processing*, vol. 25, no. 7, pp. 2971-2982, 2016.

[3] J. M. Guo and S. Sankarasrinivasan. "Digital Halftone Database (DHD): A Comprehensive Analysis on Halftone Type," *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1091-1099, 2018.

[4] M. Mese and P. P. Vaidyanathan, "Optimized halftoning using dot diffusion and methods for inverse halftoning," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 691-709, 2000.

[5] L. W Chang, and J. R. Liao, "Noise shaping for direct binary search image halftoning," *Journal of Electronic Imaging*, vol. 29, no. 3, 2020.

[6] D. L. Lau, R. Ulichney, and R. AGonzalo, "Blue and green noise halftoning models," *IEEE Signal Processing Magazine*, vol. 20, no. 4, pp. 28-38, 2003.

[7] Isola, Phillip et. al., "Image-to-image translation with conditional adversarial networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125-1134. 2017.

[8] Z. Wang., et. al., "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing,* vol. 3, no. 4 pp. 600-612, 2014.