H-GAN: Deep Learning Model for Halftoning and its Reconstruction

Jing-Ming Guo, Senior Member, IEEE, and Sankarasrinivasan S
Dept. of Electrical Engineering, National Taiwan University of Science and Technology, Taiwan
jimguo@seed.net.tw, D10507805@mail.ntust.edu.tw

Abstract—Digital halftone deals with transforming a gray scale image into its printable binary version. In this paper, a generative adversarial network-based model is proposed to perform both halftoning and its structural reconstruction. The GAN model is based on the concept of unpaired image to image translation and the model learns both transformations simultaneously. For optimal training, the model is feed with high structural data and the model architecture is also modified in accordance to this optimal training. From results, it has been validated that the model performs with consistent accuracy in both transformations.

I. INTRODUCTION

Digital halftoning [1] is a method to transform the continuous tone gray scale image into its appropriate binary image which is generally used for printing. When the printed binary pattern is viewed by the humans from a specific distance, it is perceived as continuous tone image due to the low pass nature of the human eye. The technique is widely adopted in personal and commercial printing, and the printing engines such as laser and ink-jet printers are predominantly used.

Fig. 1. Digital Halftoning-Dispersed Dot

Halftoning patterns is mainly characterized based on the distribution of black and white dots. For a given input image, there are many possibilities to generate an approximate binary patterns. Among them, it is very challenging to learn and generate the binary patterns corresponding to the best visual quality. From human perception viewpoint, the binary pattern pertains to the blue-noise distribution is perceived as the visually pleasant. The blue-noise binary patterns are often characterized through dispersed patterns which comprise of homogeneous distribution of black and white dots. In this paper, a blue noise based dithering screen is used and the halftone images is constructed using the ordered dithering process. The inverse halftoning can be seen as an image reconstruction problem in which the model has to recover the gray scale information from the approximated binary values. The main challenge exist in the inverse halftoning lies in accurate reconstruction of structural information.

In literature, the halftoning is generally performed using the mathematical techniques and does not involve any learning. Hence this approach is novel for halftone generation based on the deep learning methods. On the other hand, for inverse halftoning, many conventional mathematical approaches and deep learning methods are utilized mostly in the recent times. Most of conventional approaches involves look up tables-based approach and is trained on limited images [3], [4]. The methods also possess huge memory requirement and involves huge time consumption to search and obtain the best reconstruction patch. Recent methods are mainly based on the U-Net architecture using pre-trained CNN and has limited performance towards structural reconstruction. In comparison with others, the proposed model can effectively perform both halftoning and inverse halftoning. The contribution of the methods are as follows:

I) Development of dispersed dot halftone database to obtain superior quality halftone images in the forward transformation

II) Simplified and optimized image to image translation model to perform halftoning and inverse halftoning.

III) Improved training using structural patches which can tremendously reduce training time and increase efficiency.

II. H-GAN METHODOLOGY

The proposed GAN model to perform the halftoning and inverse halftoning is inspired from the Cycle-GAN [5]. The existing architecture is modified to suits the case of halftoning. As shown in Fig. 2, the model comprises of two transform mapping function such as $G_{AB}$ and $G_{BA}$, in which $G_{AB}: A \rightarrow B$ deals with halftoning and $G_{BA}: B \rightarrow A$ deals with inverse halftoning. On the other hand, two adversarial discriminators such as $D_A$ and $D_B$ are used two distinguish the halftone and inverse halftones images.

$$L_G(G_{AB}, D_B, A, B) = E_{B-\rho_{data}(B)} \left[ \log D_B(B) \right] +$$

$$E_{A-\rho_{data}(A)} \left[ \log(1 - D_B(A)) \right]$$

The existence of adversarial loss will help the model to learn the mapping from one domain to the other, but this does not guarantee a unique output, and hence an additional cycle consistency (CC) loss is integrated to handle this issue. The CC loss can be formulated as

$$A \rightarrow G(A) \rightarrow F(G(A)) \sim A$$

The consolidated loss function is provided as

$$L_T = L_G(G_{AB}, D_B, A, B) + L_G(G_{BA}, D_A, A, B) + \lambda L_{CYC}(G_{AB}, G_{BA})$$
During training, the $\lambda$ is set to 20 which controls the importance of two objectives. The Adam optimizer is used with a batch size of 1 and the learning rate is fixed at 0.001.

The network architecture is as follows:

**Generator:**
- $c7$-$32$-$s1$, $d128$, $R256$, $R256$, $u64$, $c7$-$32$-$s1$; where $c7$-$32$ corresponds to 7x7 convolution with instant normalization and ReLu layer, 32 indicate number of filter and s1 refers to stride 1. Further $d128$ denotes a smaller convolution filter size of 3x3 with 128 filters and for all ‘d’ layers stride 2 is maintained. $R256$ denote residual blocks that contains two 3x3 convolutional layers and finally $u64$ denotes 3x3 convolution with fractional stride of 0.5. Three residual layers are removed from existing model and number of filters in the first and last layer is reduced half than original version.

**Discriminator:**
- The PatchGAN model is adopted for the discriminator and consist of: $C64-C128-C256-C512$. The training is performed for the image patch resolution of 128x128. From training results, it has been inferred that the model trained with halftone patches from the whole image does not perform well and can be significantly improved by adjusting the training through structural patches. For instance, the structural patch of a sample training image is shown in Fig. 3. It can be seen that non-structural patches do not contain any useful information for structural reconstruction and mostly corresponds to smooth regions which are quite easy to reconstruct.

![Fig. 3. Structural patches from a halftone image](image)

For training, 10,000 random patches are extracted from the Place365 and DHD database [2] and ordered dithering (for dispersed dot) is used for database construction. The ratio of 80:20 is maintained between structural and non-structural patches respectively. For testing, 1000 patches are used and the result are provided in Table I.

<table>
<thead>
<tr>
<th>Reconstruction</th>
<th>DHD</th>
<th>Place365</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context-Driven</td>
<td>26.55</td>
<td>28.66</td>
</tr>
<tr>
<td>Proposed Method</td>
<td><strong>26.95</strong></td>
<td><strong>29.06</strong></td>
</tr>
</tbody>
</table>

![Fig. 4. Halftone Reconstruction](image)

The results are the images generated in the reverse cycle $G_{BA}$, and from the Fig. 4, it can be easily seen that the proposed method has better image quality in terms of Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). In the forward cycle $G_{AB}$, the halftone image is reconstructed from the gray scale image. In addition to $G_{AB}$, a thresholder unit is also provided to ensure the pattern homogeneity. The generated halftone image achieved consistent similarity with the actual halftone image.

### III. Conclusion

In this paper, a generative model-based solution is proposed to perform halftoning and its reconstruction. The deep learning model is optimized with respect to halftoning and training is performed with structural patches. The image reconstruction of the proposed model is found to be consistent with the existing approaches. The critical advantage of the proposed model is, it can perform both halftoning and its inversion, whereas the existing model is limited to one side transformation.

### REFERENCES