UDBNET: Unsupervised Document Binarization Network *via* Adversarial Game

Amandeep Kumar[a*], Shuvozit Ghose[b*], Pinaki Nath Chowdhury[c], Partha Pratim Roy[d], Umapada Pal[c]

(*Equal Contribution) [a]Techno Main Salt Lake, Sector V, Kolkata, India. [b]Institute of Engineering and Management, Kolkata India.

[c]Indian Statistical Institute, Kolkata, India. [d]Indian Institute of Technology, Roorkee, India.

1. Problem

Degraded document image binarization is very challenging tasks due to various reasons. Although state-of-the-art binarization methods works for supervised setup, [1] first introduced unsupervised setup for document image binarization. It consists of Texture Augmentation Network (TANet) that superimposes noisy appearance of degraded document on clean binary image to generate multiple degraded image of same textual content with various noisy textures. Next Binarization Network (BiNet) is used to get back the clean version of the document image.

This method has several limitations. (i) TANet is completely unaware about the content at which it is conditioned on. Thus, the corresponding discriminator can not verify if the content of the generated noisy image remain consistent or not. (ii) there exist no performance quantifier that validates the performance of the BiNet on real degraded noisy image. (iii) the Binarization Network (BiNet) has dataset bias towards generated noisy images. But, to adddress the dataset bias, BiNet does not use any kind of formulation.

2. Motivation

In our observation, these limitations are due to the fact that the TANet and BiNet both employ straight-forward two-player Generative Adversarial Network (GAN) objectives and model two different uncorrelated conditional distributions. We address these limitations by introducing adversarial minmax game in the domain of unsupervised document image binarization.

5. Conclusion and Future Work

In this paper, we have proposed a novel approach towards document binarization by introducing three-player min-max adversarial game. We introduce a joint discriminator which tries to couple the Adversarial Texture Augmentation Network (ATANet) and Unsupervised Document Binarization Network (UDBNet) so that it can tackle the dataset bias problem and perform well on the real degraded document image. The proposed framework is simple and easy to implement. We demonstrate the effectiveness of our system by conducting experiments on publicly available DIBCO datasets. The results of the experiment show the superiority of our proposed model.

References

[1] A. K. B. et al., "Improving document binarization via adversarial noise-texture augmentation," in *ICIP*, 2019.

[2] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE-TSMC*, 1979.

[3] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern recognition*, 2000.

[4] N. Howe, "Document binarization with automatic parameter tuning," *IJDAR*, 2013.

[5] B. S. et al., "Binarization of historical document images using the local maximum and minimum," in *DAS*, 2010.

[6] F. J. et al., "Degraded document image binarization using structural symmetry of strokes," *Pattern Recognition*, 2018.

[7] G. D. Vo and C. Park, "Robust regression for image binarization under heavy noise and nonuniform background," *Pattern Recognition*, 2018.

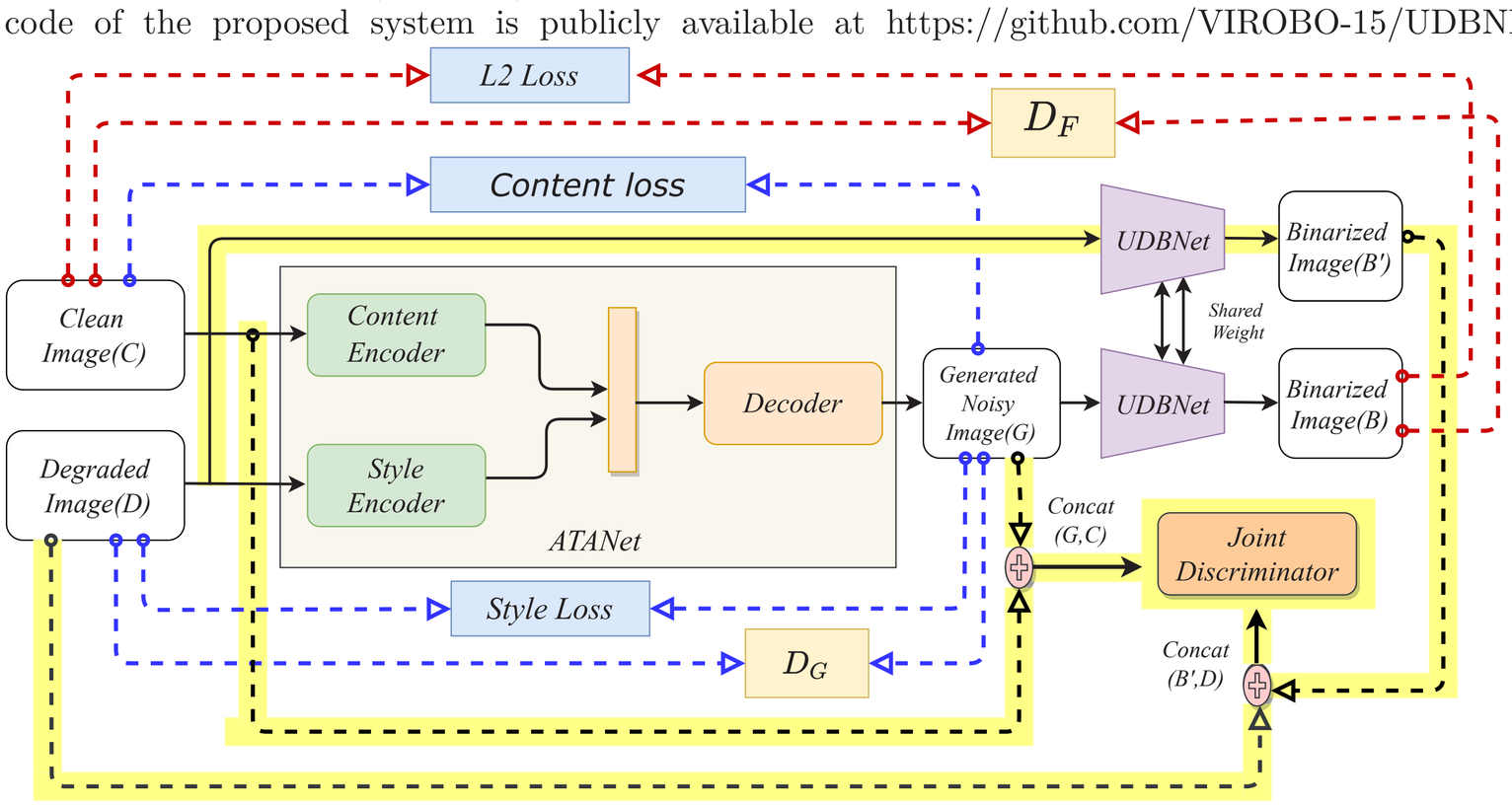
[8] Q. V. et al., "Binarization of degraded document images based on hierarchical deep supervised network," *Pattern Recognition*, 2018.

[9] F. W. et al., "Document image binarization using recurrent neural networks," in *DAS*, 2018.

[10] S. He and L. Schomaker, "Deepotsu: Document enhancement and binarization using iterative deep learning," *Pattern Recognition*, 2019.

3. Proposed Methodology

In this paper, we address the aforementioned limitations by introducing adversarial minmax game in the domain of unsupervised document image binarization. Similar to the TANet and BiNet, we propose Adversarial Texture Augmentation Network (ATANet) and Unsupervised Documenet Binarization Network (UDBNet) which utilize three-player GAN objectives. The proposed third player is a joint discriminator tries to couple both the Adversarial Texture Augmentation Network (ATANet) and Unsupervised Document Binarization Network (UDBNet). Our three-player min-max adversarial game comes to an end, when the distribution modelled by the Adversarial Texture Augmentation Network (ATANet) and the Unsupervised Document Binarization Network (UDBNet) align to the same joint distribution over time. The source code of the proposed system is publicly available at https://github.com/VIROBO-15/UDBNET.



4. Results

Comparsion of Our method with Baseline Methods

Methods	F-Measure	F_{PS}	PSNR	DRD
UDBNet-CL	92.7	95.8	19.9	2.6
UDBNet-GRL	93.2	96.0	20.1	2.4
Ours	**93.4**	**96.2**	**20.1**	**2.2**

Quantative results on H-DIBCO 2016 and DIBCO 2011 dataset

Methods	H-DIBCO 2016 Dataset				DIBCO 2011 Dataset			
	F-Measure	F_{PS}	PSNR	DRD	F-Measure	F_{PS}	PSNR	DRD
Otsu [2]	86.6	89.9	17.8	5.6	82.1	84.8	15.7	9.0
Sauvola [3]	84.6	88.4	17.1	6.3	82.1	87.7	15.6	8.5
Howe [4]	87.5	92.3	18.1	5.4	91.7	92.0	19.3	3.4
Su [5]	84.8	88.9	17.6	5.6	87.8	90.0	17.6	4.8
Jia [6]	90.5	93.3	19.3	3.9	91.9	95.1	19.0	2.6
Vo [7]	87.3	90.5	17.5	4.4	88.2	90.3	20.1	2.9
Vo [8]	90.1	93.6	19.0	3.5	93.3	96.4	20.1	2.0
Westphal [9]	88.8	92.5	18.4	3.9	-	-	-	-
DeepOtsu [10]	91.4	94.3	19.6	2.9	93.4	95.8	19.9	1.9
Bhunia [1]	92.3	95.4	19.9	2.7	93.7	96.8	20.1	1.8
Ours	**93.4**	**96.2**	**20.1**	**2.2**	**95.2**	**97.9**	**20.4**	**1.5**