Unsupervised Document Binarization of Engineering Drawings via Multi Noise CycleGAN

Luqman Hakim Rosli, Yew Kwang Hooi, Ong Kai Bin
Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Malaysia

***Abstract*—The task of document binarization of degraded complex documents is tremendously challenging due to the various forms of noise often present in these documents. While the current state-of-the-art deep learning approaches are capable for the removal of various noise types in documents with high accuracy, they employ a supervised learning scheme which requires matching clean and noisy document image pairs which are difficult and costly to obtain for complex documents such as engineering drawings. In this paper, we propose our method for document binarization of engineering drawings using 'Multi Noise CycleGAN'. The method utilizing unsupervised learning using adversarial and cycle-consistency loss is trained on unpaired noisy document images of various noise and image conditions. Experimental results for the removal of various noise types demonstrated that the method is able to reliably produce a clean image for any given noisy image and in certain noisy images achieve significant improvements over existing methods.**

Keywords*—*Image processing and computer vision; generative adversarial networks; document binarization; deep learning

I. Introduction

In the current Industry Revolution 4.0, images and documents are important methods to acquire information. Specifically, this information is in the form of texts contained within images and documents. Document binarization is a crucial pre-processing step when attempting to analyse and extract the information. The goal of document binarization is to retrieve and extract underlying text from the background of documents. More accurate binarization methods allow for performance improvement of document analysis such as text line and word segmentation and Optical Character Recognition (OCR). Successful and accurate OCR of documents allows for better indexing and searching of documents within document management systems.

Among the many challenges with document images is that they are normally affected by various types of noise and environmental degradation. Examples are uneven illumination, ink or page stain, background colour, shadows, blue and white noise. These noises can come from the state of the document themselves when scanned or bad conditions of digitization. Text extraction from these documents with noisy backgrounds especially those with randomized noise poses a tricky challenge as these noises can obstruct text and possibly not picked up by denoising methods.

In recent years, considerable amount of research pertaining to document image enhancement and recognition has been studied and proposed. Document binarization is an active and ongoing research with steady paper submissions on a yearly basis to Document Image Binarization Competition (DIBCO) which is held by International Conference on Document Analysis and Recognition (ICDAR).

Before the widespread usage of deep learning, popular conventional methods for document binarization are global threshold methods, local threshold methods, statistical-based methods and edge-based. Conventional methods typically focus on a specific noise such as bleed through [1], [2]. The limitation of conventional methods is a reduction in noise removal performance if the document contains various and severe noise and degradation such as a combination of shadow, bleed through and uneven illumination.

Recently deep learning-based document binarization methods have shown significant performance improvement when compared to conventional methods. Built upon the framework Generative Adversarial Networks (GANs), proposed by Goodfellow et al. [12], supervised GANs have been utilized and proven capable to address more challenging noise and multiple degradation and noise. Yun Hsuan Lin [4] uses conditional GANs to remove shadows from complex background colour in documents. Liu et al. [5] utilizes a combination of Recurrent Neural Network with conditional GAN to address complex background noise, colour, and watermarks. Deep learning methods such supervised GAN's are trained in a supervised manner hence matching noisy documents and clean target document pairs are required which are difficult to obtain or generate. Another limitation of supervised GANs and other deep learning methods is that the methods are trained to remove specific noise type which often leads to poor results if the paired training data is not sufficiently many and does not match the noise type.

The ability to denoise and preserve original text and graphics is crucial for accurate OCR and digital document information management. Complex documents such as technical drawings often contain various forms of noise and degradation. Coupled with their irregular text and graphics greatly reduces the performance of an OCR system. The combination of randomized noise and irregular text and graphics results in paired training data for supervised denoising to be difficult or near impossible to obtain. Furthermore, with both the noise and text and graphics in technical drawings being very different from one and another, this results in supervised solutions to be trained multiple times for different combinations of noise, text, and graphics.

Zhu et al. [11] proposed a Cycle-Consistent Adversarial Network (CycleGAN), an unsupervised image to image translation model which does not require paired training data and provides solution to the problems of limited paired datasets. Additionally, CycleGAN has been demonstrated as a capable and flexible denoising method for images with complex noise by Song [14] in denoising complex noisy satellite images into clean satellite images.

Inspired by the framework and capability of CycleGAN, this paper proposes a framework of a modified cycle consistency generative adversarial networks, dubbed 'Multi Noise CycleGAN' for flexible engineering drawing document binarization catering to randomized noise.

The targeted contributions of this paper can be summarized as follows.

- The proposed method focuses on the multi degradation and noise with limited training data availability.
- This framework could be exploited and adjusted to address any complex document degradation problem.
- The proposed hybrid method focuses on addressing current industrial challenge, digitization of historical engineering drawing.

The rest of this paper is organized as follows: Section II reviews associated literature and related works on document binarization and document noise removal. Section III presents the proposed Multi Noise CycleGAN, and the modifications made on the CycleGAN. Section IV presents the experimental results, performance analysis and comparison against other solutions. Section V concludes the paper.

II. RELATED WORK

A. Document Binarization

Research studies covering document image binarization have been conducted and reported in the literature over the past decade. Document Binarization is the process of discarding unnecessary noisy information from documents to preserve meaningful text for text recognition and text extraction. Text in scanned documents is more dense than natural images and contain more contextual information. Emphasis is placed to ensure the text-extracted are similar to the original source. Document image binarization methods are typically categorized into two types: traditional binarization algorithms and deep learning semantic segmentation algorithms. Thresholding is the most basic commonly used method, which sets pixels below a threshold value to 0 and the rest to 1. Global, local and hybrid are among the main types of thresholding. Otsu algorithm [6] is widely regarded as among the most used classical global binarization method.

To drive the progress of document binarization, Document Image Binarization Content (DIBCO) was introduced. Document Image Binarization Content (DIBCO) is a both a document database meant to challenge and serve as a general benchmark for researchers in the field of document binarization with a new document database provided by a specialist every year and a contest of binarization. Chen [7] highlights deep learning models have become the state-of-the-art in document binarization where in 2017, all top six submissions to DIBCO have utilized deep learning model. The document database contains several degradation types that commonly occur in historical documents [7], ranging from stains, bleed through ink, stamps, multiple colour text and border noise.

B. Supervised Deep Learning Document Denoising

Active studies on supervised deep learning document denoising focuses on further improving the accuracy of document binarization alongside covering more degradation and noises within a single document. Souibgui [8] proposes Document Enhancement Generative Adversarial Networks (DE-GAN) to clean and restore severely degraded document images. The DE-GAN reported successful and accurate in addressing blurriness, watermark and ink stain among other noises as well. Most recently, Suh [9] proposes a two stage GAN architecture, addressing document binarization with coloured background, watermark and ink bleed achieving higher results than most state-of-the-art solutions. The architecture proved successful and accurate for shipping labels as well. Both methods are tested on DIBCO datasets for standardization; results at the time of paper publish showcased improvement over other state-of-the-art methods.

Furthermore, Chung [10] proposed a CNN-based binarization for historically degraded drawing maps. Over the years, there have not been many studies and research covering complex documents such as historical drawing maps. Chung [10] CNN based binarization method is an improvement over existing works covering the same document domain, binarization of complex documents consisting of yellowing noise, folded noises and preserving components. Recent works in supervised deep learning document denoising showcases the maturity of methods and algorithms in tackling standardized document binarization test. However, few works address document binarization for complex documents.

C. Cycle consistency Generative Adversarial Network (CycleGAN)

To eliminate the reliance on paired data for training, several unsupervised learning-based methods have been developed for image enhancement inclusive of image denoising. Among them, is a general-purpose method, unsupervised image to image to translation model, cycle consistency generative adversarial network (CycleGAN) proposed by Zhu [11]. CycleGAN allows for the generated image to be as close as target image without paired data and supervision. CycleGAN is derived from generative adversarial networks [12] which utilizes adversarial training, pitting two neural networks against one another. The study of GAN has gained a lot of traction and many variations have been developed to address different problems in the field of computer vision and image processing.

D. Application of GAN and CycleGAN in Image Denoising

The utility of GANs and CycleGAN as an image denoiser has been researched and tested across different industry. Xiong [13] has developed a two stage GAN for low light image enhancement.

In the medical field, due to the intense difficulty of obtaining paired images, modifications were made to cycleGAN as a means to obtain higher resolution and cleaner images to aid doctors in receiving more details from the pictures and increase the diagnostic accuracy. Khan [15] opt for optimal transport cycleGAN to reconstruct high resolution MR images while Lee [16] utilizes attention guided – β CycleGAN to remove metal artifacts from CT images. The works above shares the common goals in removing random noise while preserving the original image through unsupervised image denoising and with limited training data. In the next subsection, past works incorporating CycleGAN ability as an image denoiser at the domain of document binarization is reviewed.

E. Application of cycleGAN in Document Binarization

In the areas of document denoising and binarization, Sharma [17] has applied CycleGAN to the denoising of scanned documents focusing on the removal of watermark, blurriness and background noise removal. Utilizing the design of CycleGAN, Kumar [18] addresses the works of Bhunia [19] for text extraction from documents as the original works of Bhunia [19] requires paired training data, through the development of UDBNET architecture which adds another layer unto the architecture of Bhunia [19] and merging them with a joint discriminator to allow for accurate text extraction without paired data. Tensmeyer [20] extends the CycleGAN model to trained on ground truth binarization masks to produce realistic synthetic data for document binarization of DIBCO datasets.

From recent works, unsupervised document denoising has been tested and refined on standardized degraded documents. However, there exist a research gap in document binarization of complex documents such as technical drawing.

F. Conceptual Research Framework

Fig. 1 provides the framework for a disciplined research approach to the factors affecting OCR performance.

The conceptual framework above highlights the factors and their weightage that contributes to the overall OCR engine performance. Realistically every document requires a certain level of denoising to allow the OCR to accurately extract the underlying text. While there are four main variables affecting the denoising layer, 2 of them, noise type & amount and training data availability has a higher weightage because these are not easily controlled by the user. Commonly used metrics to evaluate OCR performance are Character Error Rate (CER) and Word Error Rate (WER) [24]. [25]. CER and WER indicate the amount of character for CER and word for WER that the applied OCR did not read or generate correctly.

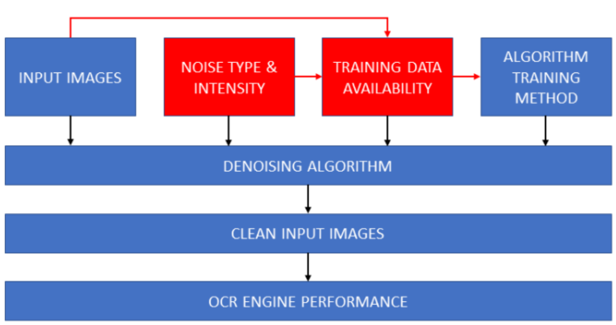


Fig. 1. Overview of OCR performance conceptual framework.

Noise in images can be present due to many factors, environmental, condition of the digital capture device and the condition of the source item. The randomness of the noise has a direct effect unto the training data availability as rarer and more complicated noise generally do not occur often and a sufficient number of data is required to effectively train the denoising algorithm.

The availability of training data is second factor with higher weightage with direct impact to the selection of algorithm training method. In an ideal scenario, even with rare noise type, so long as paired training is available, supervised deep learning denoising method can be utilized. However, that is not often the case. While more mainstream and widespread documents or images have high volumes of data available, certain images especially those in technical areas, engineering, manufacturing or medicine, paired training data are difficult and sometimes impossible to obtain.

As a summary, the critical point for good OCR engine performance lies in the training data. There are very few works that have addressed lack of training data for complex documents. Failure to successfully binarize and conduct text extraction on these documents will result in loss of information as there are enormous number of complex documents from before the digital era.

III. Proposed Method

Image denoising targets to restore a clean image from a noisy one to improve the overall quality of the image for better OCR result. The problems of engineering drawing binarization is treated as an image-to-image translation task where the objective is to produce and generate clean document images and preserve the original content from its noisy counterpart.

CycleGAN has shown to be capable in addressing image to image translation problems in situations there is a lack of paired training dataset. With CycleGAN specific capability of able to work without the constraint of one-to-one mapping between the input image and target image whilst keeping its ability to learn such image-to-image translations, has convinced us to investigate and modify the framework accordingly as a denoising layer for engineering drawings as it is extremely difficult and limited to obtain clean engineering drawings alongside its corresponding noisy counterpart. In Fig. 2 the overview the proposed framework is highlighted.

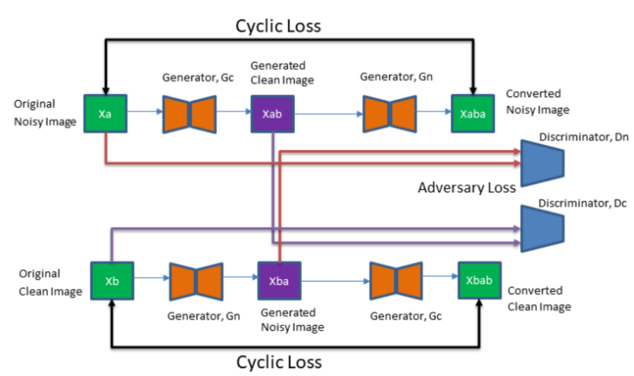


Fig. 2. Overview of CycleGAN.

CycleGAN leverages cycle-consistency loss to get around the problem of learning meaningful transformations in unpaired datasets. This means that if an image is converted from source distribution to target distribution and back, samples can be acquired from source distribution. The cycle consistency loss is incorporated via CycleGAN two generators and two discriminators.

The first generator, G_c converts an image taken from noisy domain, X_a to produce an output image in the targeted clean domain, X_{ab}. To promote and enforce effective relation between noisy and clean images, CycleGAN must be able to learn the necessary features that can be incorporated to turn back generated clean images to its original noisy counterpart. A similar process but in reverse takes place with the second generator, G_n , to convert clean images, X_b to noisy images.

The responsibility of the discriminator is to be able to identify real and fake images generated by the generators ideally as to defeat the generator via rejecting the images produced by it. The generator and discriminator engage in a competition like manner till the generator is able to produce images that are indistinguishable from the original input images. This is expressed as adversary loss.

The proposed, 'Multi Noise CycleGAN', utilizes a similar network structure to that of the original CycleGAN by Zhu [11] with modification made to the existing network which is as follows:

1) 15 ResNet blocks used to build the Generator Architecture as shown in Fig. 3.

2) Individual learning rate for Generator and Discriminator as opposed to same.

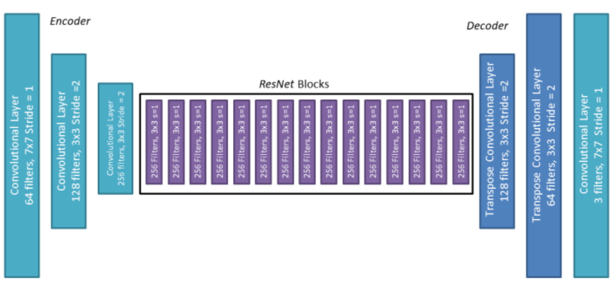


Fig. 3. Overview of modified CycleGAN generator architecture.

Full objective loss is as follows:

$$L\,(G_N, F_L, D_N, D_L) = L^{p}_{GAN,N}\,(G_N, D_N, X, Y) + L^{p}_{GAN,L}\,(F_L, D_L, Y, X) + \lambda L_{cyc}\,(G, F) + L_{IDENTITY}\,(G, F) \quad (1)$$

IV. Experiments and Results

The following subsections describe the experiments that have been conducted using 'Multi Noise CycleGAN'. The proposed dataset, training images and test images alongside the evaluation metrics are described as well.

A. Datasets' Details

In this paper, the experiments are conducted on the noisy engineering drawing dataset type proposed by Chung [10], It consists of 35,960 paired images containing noisy background, yellowed area, and folded lines and their respective clean image. The dirty images are collected and compiled from real degraded as built drawing maps. The noises ae noisy background, yellowed area, and folded lines. The dirty images contain varying degrees of noise level, image clarity and complexity of text and graphics.

For the training phase, 5,083 unpaired dirty images with 256x256 size and 5,063 clean images 256x256 size is used to train 'Multi Noise CycleGAN'.

For the testing phase, to evaluate the overall performance of unsupervised image to image translation, 3,200 pair of images that best represents the structure of engineering drawings containing engineering drawing symbols and diagrams were selected (dirty images and their corresponding clean versions) from the 35,960 paired images and does not include images used during training.

The 3,200 testing datasets are a mixture of four different noises namely as follows:

1) Noisy / Complex background.
2) Less Noisy Background.
3) Noisy Background with discolouration.
4) Noisy Background with poor illumination.

B. Evaluation Metrics

In this paper, to compare and evaluate the performance of unsupervised image denoising with other binarization method, two metrics are used. Structural Similarity Index (SSIM) and Peak Signal to Noise Ratio (PSNR).

The equation for SSIM is as follows:

$$SSIM\ (x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)((\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2)$$

Where SSIM is the Structural Similarity Index: (x,y) are respective coordinates indicating a nearby NxN window: σ_x , σ_y are the variances of intensities in x,y directions, σ_{xy} is the covariance and μ_x, μ_y are the average intensities in x,y directions.

The equation for SSIM is as follows:

$$PSNR\ (I_{original}\ , I_{clean}) = 10log_{10}\frac{255^2}{MSE} \quad (3)$$

Where MSE is defined as the Mean Square Error. $I_{original}$ is the original image and I_{clean} is the clean, denoised version. MSE is defined as follows:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I_{original}(i,j) - I_{clean}(i,j)]^2 \quad (4)$$

Where m and n respectively are the size of the input images.

For the testing phase, 3,200 pair of images outside of the training data is selected (dirty images and their corresponding clean versions) is used to evaluate the overall performance of unsupervised image to image translation.

C. Experimental Results

Fig. 4 highlights three examples when 'Multi Noise CycleGAN' method is applied unto the test dataset. From the end results, from a visual standpoint, the image is cleaned to a satisfactory level however it is challenging to differentiate a real image from a generated one when the differences between those two are considerably small. As such, quantitative comparison on the proposed method is applied using the aforementioned two metrics: Structural Similarity Index (SSIM) and Peak Signal to Noise Ratio (PSNR).

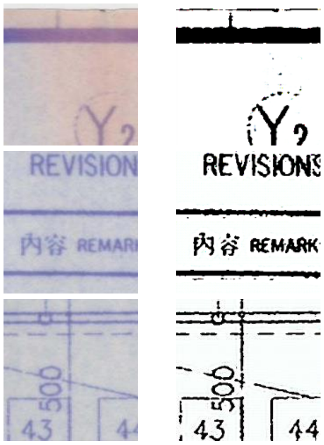


Fig. 4. Multi Noise CycleGAN. From left to right: noisy image, result of Multi Noise CycleGAN.

Guided by the two-evaluation metrics defined above, Table I as shown below summarizes the average results across 3,201 images comparing the performance of Multi Noise CycleGAN against other state-of-the-art methods.

From the results above, 'Multi Noise CycleGAN' on average has achieved PSNR of 15.828 db, and SSIM of 0.806. Therefore, it is shown that the proposed method has achieved comparable results in PSNR and SSIM against state-of-the-art methods with the added advantage of not having to use paired images for the training phase.

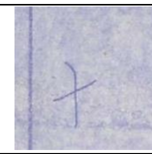
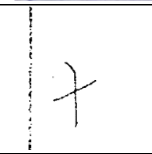
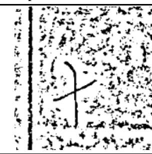
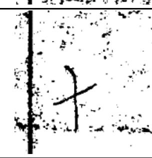
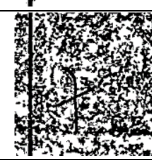
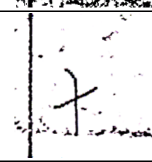
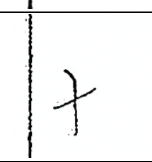
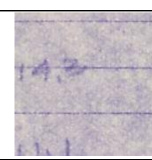
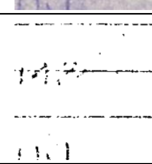
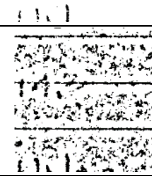
Table II highlights few test cases for 1 noise category where the proposed 'Multi Noise CycleGAN' achieves better results in terms of PSNR/SSIM against state-of-the-art method indicating an area of image denoising where it is more advantageous to utilize the proposed 'Multi Noise CycleGAN' method.

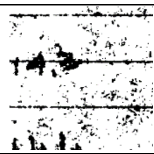
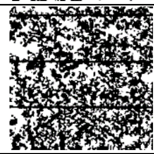
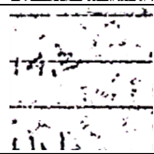
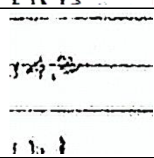
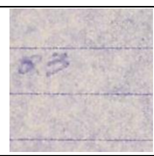
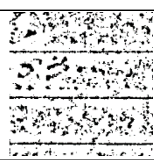
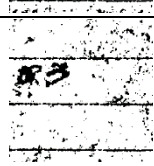
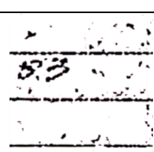
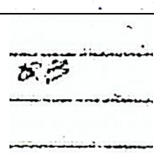
It can be observed that the proposed method from a visual perspective as shown in Table II generates a clean image whilst retaining important image details as well without any additional artifacts created.

TABLE I. QUANTITATIVE COMPARISON WITH STATE-OF-THE-ART METHODS ON THE PROPOSED NOISY IMAGE DATASET

Method	PSNR	SSIM
Sauvola [21]	15.418	0.883
Niblack [22]	8.236	0.352
Otsu [6]	16.139	0.769
Jia et al [9]	17.483	0.839
Isola et al [3]	18.798	0.869
Multi Noise CycleGAN	15.828	0.806

TABLE II. QUANTITATIVE COMPARISON WITH STATE-OF-THE-ART METHODS ON MULTIPLE TEST CASES ON NOISY / COMPLEX BACKGROUND NOISE

Test Case #1 Noisy / Complex Background			
Method	Result	PSNR	SSIM
Noisy Image		-	-
Sauvola [21]		14.464	0.908
Niblack [22]		7.777	0.287
Otsu [6]		12.822	0.691
Jia et al [23]		3.496	0.078
Isola [3]		14.931	0.830
Multi Noise CycleGAN		16.527	0.915
Test Case #2 Noisy / Complex Background			
Method	Result	PSNR	SSIM
Noisy Image		-	-
Sauvola [21]		13.638	0.855
Niblack [22]		8.954	0.390

Otsu [6]		11.716	0.602
Jia et al [23]		3.175	0.092
Isola [3]		14.834	0.790
Multi Noise CycleGAN		16.506	0.822
Test Case #3 Noisy / Complex Background			
Method	Result	PSNR	SSIM
Noisy Image		-	-
Sauvola [21]		13.725	0.856
Niblack [22]		8.930	0.391
Otsu [6]		12.288	0.652
Jia et al [23]		12.118	0.803
Isola [3]		15.446	0.809
Multi Noise CycleGAN		16.858	0.817

V. Conclusion

This paper presents a novel model that utilizes Cycle consistent adversarial networks for complex image denoising. The method, dubbed ‘Multi Noise CycleGAN’ architecture utilizes and modifies CycleGAN to the task of complex image denoising. The method is able to both denoise simple and complex noise due to the inherent property of CycleGAN of cycle-consistency. It is also noteworthy that the method uses unpaired images as training dataset, and therefore does not require ground-truth clean images.

Results show that the proposed method achieves denoising results comparable to state-of-the-art methods. The values of PSNR and SSIM obtained are at same level with state-of-the-art method found in the research literature with some test cases showcasing where the proposed method is overall better.

Future work is targeted to test the feasibility of ‘Multi Noise CycleGAN’ in tackling additional document binarization problems such as blur and watermark. Improvements to the architecture, to accommodate the addition of blur and watermark noises will allow for even better flexibility of noise removal, improving the accuracy of document binarization of engineering.

Acknowledgment

This research/paper was fully supported by Universiti Teknologi PETRONAS, under the Yayasan Universiti Teknologi PETRONAS (YUTP) Fundamental Research Grant Scheme (YUTP-PRG/015PBC-005).

References

[1] M. R. Yagoubi, A. Serir, and A. Beghdadi, “A new automatic framework for document image enhancement process based on anisotropic diffusion,” in 2015 13th International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2015, pp. 1126–1130.

[2] B. Sun, S. Li, X.-P. Zhang, and J. Sun, “Blind bleed-through removal for scanned historical document image with conditional random fields,” IEEE Transactions on Image Processing, vol. 25, no. 12, pp. 5702–5712, 2016.

[3] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

[4] Lin, Y. H., Chen, W. C., & Chuang, Y. Y. (2020). BEDSR-Net: A Deep Shadow Removal Network From a Single Document Image. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 12905-12914).

[5] Liu, S., Zhang, F., Chen, M., Xie, Y., He, P., & Shao, J. (2019). Document Binarization using Recurrent Attention Generative Model. In BMVC (p. 95).

[6] N. Otsu, “A threshold selection method from gray-level histograms,” IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1, pp. 62–66, 1979.

[7] Chen, X., Jin, L., Zhu, Y., Luo, C., & Wang, T. (2021). Text recognition in the wild: A survey. *ACM Computing Surveys (CSUR)*, *54*(2), 1-35.

[8] Souibgui, M. A., & Kessentini, Y. (2020). De-gan: A conditional generative adversarial network for document enhancement. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[9] Suh, S., Kim, J., Lukowicz, P., & Lee, Y. O. (2020). Two-Stage Generative Adversarial Networks for Document Image Binarization with Color Noise and Background Removal. arXiv preprint arXiv:2010.10103.

[10] Chung, K. L., & Hsieh, D. W. (2020). Novel and Effective CNN-Based Binarization for Historically Degraded As-built Drawing Maps. arXiv preprint arXiv:2009.05252.

[11] Zhu, J., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycleconsistent adversarial networks. CoRR abs/1703.10593 (2017), http://arxiv.org/abs/1703.10593.

[12] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

[13] Xiong, W., Liu, D., Shen, X., Fang, C., & Luo, J. (2020). Unsupervised Real-world Low-light Image Enhancement with Decoupled Networks. arXiv preprint arXiv:2005.02818.

[14] Song, Joonyoung & Jeong, Jae-Heon & Park, Dae-Soon & Kim, Hyun-Ho & Seo, Doo-Chun & Ye, Jong. (2020). Unsupervised Denoising for Satellite Imagery using Wavelet Subband CycleGAN.

[15] Khan, S., Huh, J., & Ye, J. C. (2020). Pushing the Limit of Unsupervised Learning for Ultrasound Image Artifact Removal. arXiv preprint arXiv:2006.14773.

[16] Lee, J., Gu, J., & Ye, J. C. (2020). Unsupervised CT Metal Artifact Learning using Attention-guided beta-CycleGAN. arXiv preprint arXiv:2007.03480.

[17] Sharma, Monika & Verma, Abhishek & Vig, Lovekesh. (2019). Learning to Clean: A GAN Perspective. 10.1007/978-3-030-21074-8_14.

[18] Kumar, A., Ghose, S., Chowdhury, P. N., Roy, P. P., & Pal, U. (2020). UDBNET: Unsupervised Document Binarization Network via Adversarial Game. arXiv preprint arXiv:2007.07075.

[19] Bhunia, A. K., Bhunia, A. K., Sain, A., & Roy, P. P. (2019, September). Improving document binarization via adversarial noise-texture augmentation. In 2019 IEEE International Conference on Image Processing (ICIP) (pp. 2721-2725). IEEE.

[20] Tensmeyer, C., Brodie, M., Saunders, D., & Martinez, T. (2019, September). Generating realistic binarization data with generative adversarial networks. In 2019 International Conference on Document Analysis and Recognition (ICDAR) (pp. 172-177). IEEE.

[21] J. Sauvola and M. Pietikinen. Adaptive document image binarization. Pattern Recognition, vol. 33, no. 2, pp. 225-236, 2000.

[22] W. Niblack. An introduction to digital image processing. PrenticeHall, Englewood Cliffs, NJ, pp. 115-116, 1986.

[23] F. Jia, C. Shi, K. He, C. Wang, and B. Xia. Degraded document image binarization using structural symmetry of strokes. Pattern Recognition, vol. 74, no. 2, pp. 225-240, 2018.

[24] Sporici, D., Cușnir, E., & Boiangiu, C. A. (2020). Improving the accuracy of Tesseract 4.0 OCR engine using convolution-based preprocessing. Symmetry, 12(5), 715.

[25] Alghyaline, S. (2023). Arabic Optical Character Recognition: A Review. CMES-Computer Modeling in Engineering & Sciences, 135(3).