

SHARP-GAN: SHARPNESS LOSS REGULARIZED GAN FOR HISTOPATHOLOGY IMAGE SYNTHESIS

Sujata Butte*, Haotian Wang*, Min Xian†, Aleksandar Vakanski

University of Idaho, Idaho, USA

ABSTRACT

Existing deep learning-based approaches for histopathology image analysis require large annotated training sets to achieve good performance; but annotating histopathology images is slow and resource-intensive. Conditional generative adversarial networks have been applied to generate synthetic histopathology images to alleviate this issue, but current approaches fail to generate clear contours for overlapped and touching nuclei. In this study, We propose a sharpness loss regularized generative adversarial network to synthesize realistic histopathology images. The proposed network uses normalized nucleus distance map rather than the binary mask to encode nuclei contour information. The proposed sharpness loss enhances the contrast of nuclei contour pixels. The proposed method is evaluated using four image quality metrics and segmentation results on two public datasets. Both quantitative and qualitative results demonstrate that the proposed approach can generate realistic histopathology images with clear nuclei contours.

Index Terms— Histopathology image synthesis, GAN, Nuclei segmentation

1. INTRODUCTION

Nuclei analysis is a critical step in computational pathology for quantitatively cancer grading. Most current nuclei segmentation approaches use deep learning-based end-to-end framework to map histopathology images and dense predictions [1, 2, 3, 4, 5], and requires large accurately labeled training sets. However, the manual annotation of nuclei is time-consuming, labor-intensive and expensive. Only experienced pathologists who have been trained for years can annotate the image accurately; and, on an average, it costs about two hours for an expert to annotate an image patch with 600 nuclei [6].

To overcome the challenge, Zhou *et al.* [7] proposed to re-distribute segmented nuclei to generate more nuclei masks and histopathology images. Those masks may not be the true masks of the original nuclei, and no new nuclei shapes were generated. Hou *et al.* [6] proposed a GAN network that

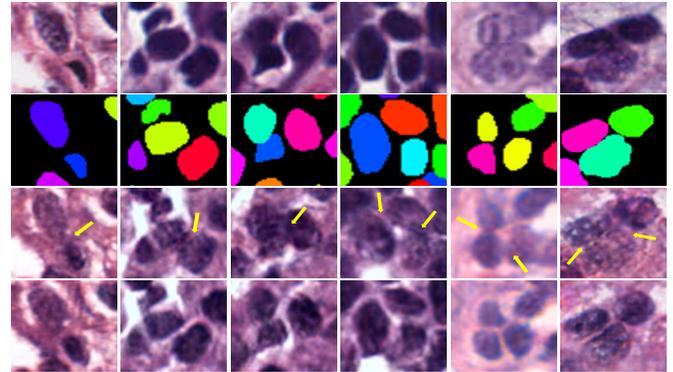


Fig. 1: First row: Real image patches. Second row: nuclei masks, different colors represent different nuclei. Third row: synthesized image patches from pix2pix-GAN [8]. Forth row: our synthesized image patches

learned the texture and shape information from the real nuclei images, and synthesized realistic histopathology images from binary image masks. However, few studies proposed an image-to-image translation network that converts the nuclei labels to images in histopathology analysis.

Conditional image synthesis has been studied widely in many fields, e.g, class-conditional models learn to synthesis images using category labels [9], image-to-image translation [8, 10]. Comparing to traditional image synthesis methods [11], deep Learning-based methods generated more realistic images. However, existing deep approaches still have limits in synthesizing histopathology images with clear nuclei boundaries. As shown Fig. 1, the approach in [8] cannot generate clear boundaries between closely clustered nuclei.

To solve this issue, we proposed a sharpness loss regularized GAN, namely Sharp-GAN to generate realistic histopathology images with clear nuclei contours. The proposed Sharp-GAN synthesizes histopathology images using distance maps of nuclei which provides full contours of nuclei. The conventional binary masks cannot accurately define the shared contours of touching nuclei. We propose the sharpness loss to reveal the contours between touching nuclei by enhancing the contrast of the contour pixels. The sharpness loss defines large penalty for contours points with small contrast.

* First authors contributed equally.

† Corresponding author.

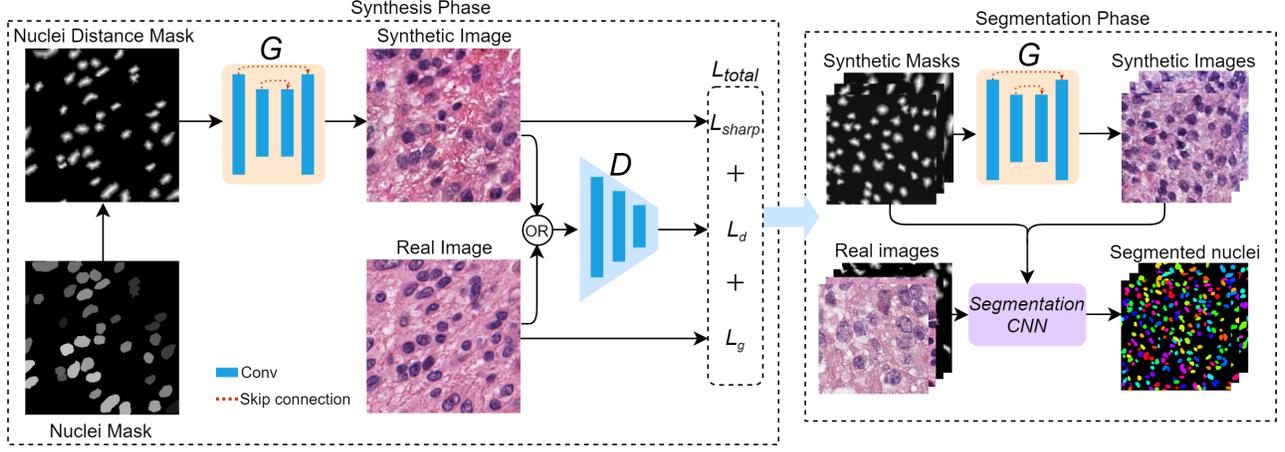


Fig. 2: Overview of the proposed Sharp-GAN.

2. PROPOSED METHOD

The generator \mathcal{G} in the proposed network learns a mapping from observed nuclei distance map $x \in \mathbb{R}_{\geq 0}^{mn}$ to realistic histopathology image y that can 'fool' the discriminator \mathcal{D} , where m and n is the width and height of the image, respectively. $x_{i,j}$ in the distance map is defined as the Euclidean distance from (i, j) to a nucleus's centroid. The discriminator \mathcal{D} is trained to detect 'fake' images produced by \mathcal{G} . The proposed network is shown in Fig. 2.

2.1. Loss function

Conditional GANs [8] tends to produce blur nucleus boundaries, and have difficulties in distinguish closely clustered nuclei. We proposed the sharpness loss to enhance the contrast of nuclei contour pixels, and the total loss is defined by

$$\mathcal{L}(\mathcal{G}, \mathcal{D}) = \mathbb{E}_{x,y}[L_1] + \mathbb{E}_x[L_2 + \beta L_{sharp}] \quad (1)$$

where

$$L_1 = \log \mathcal{D}(x, y), \quad (2)$$

$$L_2 = \log(1 - \mathcal{D}(x, \mathcal{G}(x))) \quad (3)$$

The goal is to obtain an optimal generator \mathcal{G}^* that minimizes the total loss \mathcal{L} against an optimal \mathcal{D}^* that maximizes the total loss. L_{sharp} denoted the proposed sharpness loss, the parameter β controls contribution of the sharpness loss. The generator network follows the U-Net-based pixel2pixel GAN [8]. Inspired by Graph cuts approach[12], we proposed the sharpness loss to enhance the contrast of contour pixels of nuclei (rows 3&4 in Fig. 1). The contour pixels are obtained by multiplying a binary contour map with the synthetic image. The binary contour map is generated by using the binary nuclei masks. Let c be a contour map, and \hat{y} be an synthesized

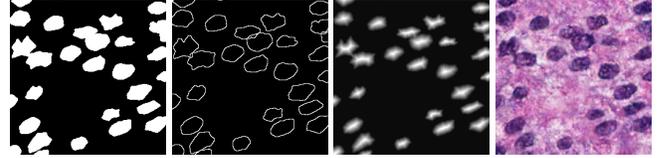


Fig. 3: A sample of (a) binary mask, (b) contour map, (c) distance map, and (d) synthetic image patch.

image of $\mathcal{G}(x)$ and the sharpness loss is defined by

$$L_{sharp}(c, \hat{y}) = \frac{1}{mn} \sum_{i,j,c_{i,j} \neq 0} S(i, j) \quad (4)$$

where $S(i, j)$ is the discrete sharpness at position (i, j) , and is given by

$$S(i, j) = \sum_{(p,q) \in N_{i,j}} \exp\left(\frac{-(g_{i,j} - g_{p,q})^2}{2\lambda^2}\right) \cdot \frac{1}{dist} \quad (5)$$

In Eq. (5), g is a gray scale image transformed from \hat{y} . $g_{i,j}$ and $g_{p,q}$ are the intensities of pixels at position (i, j) and (p, q) , respectively. $N_{i,j}$ is the set of the neighboring positions of (i, j) . λ is a threshold to determine the degrees of contrast. Large λ leads to contours with higher contrast. $dist$ is the Manhattan Distance between two contour pixels. The proposed loss gives high penalty for continues pixels when $|g_{i,j} - g_{p,q}| < \lambda$; and it defines small penalty when $|g_{i,j} - g_{p,q}| > \lambda$. We apply the eight nearest neighbors system to define neighbors.

2.2. Nucleus distant map

Many existing approaches [8] used binary image masks to input conditional GAN to generate synthesis images. However, in the histopathology image, many nuclei are closely clustered, and the binary masks cannot show the boundary between two touching nuclei. If we use the binary masks to

train the generator, the generator cannot know the accurate boundaries of nuclei and produce synthetic images with blur boundaries. To address this issue, we proposed to generate a distance map x from binary masks to clearly separate touching nuclei.

Let b be a binary mask of an image. If $b(i, j)$ is 0, and the value of the distance map at position $(i, j) : x(i, j)$ is 0; otherwise, $x(i, j)$ will be the the Euclidean distance from (i, j) to the centroid of the nucleus that contains (i, j) . Fig. 3(d) shows a sample of nuclei distance mask.

We employed the nuclei mask generation method in [6]. The method randomly generates nucleus-like polygons with variable size and irregularities. The generated polygons can be set to overlap and touch to each other. Fig. 3(a) shows a sample synthetic image with nucleus-like polygons.

3. EXPERIMENTAL RESULTS

3.1. Datasets, metrics and setting

Dataset. We evaluate the performance of the proposed approach using two public histopathology datasets, MICCAI 2015 Digital Pathology challenge (CPM-15) and MICCAI 2017 Digital Pathology challenge (CPM-17) [13]. CPM-15 contains 15 histopathology images with a total number of 2905 nuclei. CPM-17 contains 32 histopathology images with a total number of 7570 nuclei. The magnifications of these two datasets contain both $20\times$ and $40\times$. We combine two datasets in our experiments.

Evaluation metrics. Four image quality assessment metric, structural similarity index (SSIM) [14], feature similarity index (FSIM)[15], gradient magnitude similarity deviation (GMSD) [16], normalized root mean square error (NRMSE) are used to evaluate the quality of synthesized images.

We employed four metrics, detection quality (DQ) [17], segmentation quality (SQ) [17], panoptic quality (PQ) [17], aggregated jaccard index (AJI) [4] to evaluate the performance of segmentation approaches trained using different datasets. These four metrics have been applied in the nuclei segmentation [2, 3].

Training process. We feed images with size of 256×256 into the proposed network. We used random crop, flip, rotation for data augmentation. We set the batch size to 16, the initial learning rate to 0.0001, and epochs to 500 for model training.

3.2. Image quality assessment

To validate the effectiveness of the proposed distance map, we compare the proposed network using binary mask input (GAN-binary) with the proposed network with the distance map (GAN-DIST) input. In the experiments, GAN-binary model is trained using the CPM15&17 training set with nuclei binary masks, and GAN-DIST is trained using distance maps. Four image quality metrics, SSIM, FSIM, GMSD,

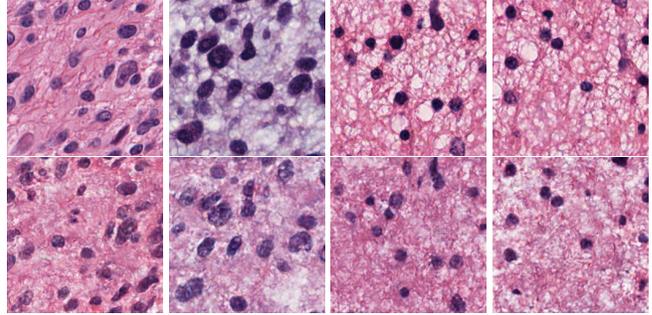


Fig. 4: Samples of original histopathology patches (first row) with the corresponding synthetic patches generated by our methods (second row).

Table 1: Image quality assessment using CPM15&17 dataset.

Model	SSIM \uparrow	FSIM \uparrow	GMSD \downarrow	NRMSE \downarrow
GAN-Binary	0.577	0.735	0.178	0.306
GAN-DIST	0.681	0.763	0.163	0.259
Sharp-GAN	0.756	0.793	0.147	0.230

and NRMSE, are employed to quantitatively evaluate the quality of synthetic images. Table 1 shows the image quality measurements of all test images. We note that GAN-DIST outperforms GAN-Binary in all metrics, i.e., the SSIM, FSIM, GMSD, NRMSE results improved 18.0%, 3.8%, 9.2%, 18.1%, respectively. The results demonstrate that using distance maps as input improves the image quality of synthetic images.

To demonstrate the effectiveness of the proposed sharpness term, we compared our model with the sharpness term (Sharp-GAN) and without the sharpness term (GAN-DIST). The two models use distance maps as input. As shown in Table 1, Sharp-GAN outperforms the GAN-DIST in all four metrics. SSIM, FSIM, GMSD, NRMSE of Sharp-GAN are 11.2%, 3.9%, 10.8%, 11.1% higher than those of the GAN-DIST, respectively.

Fig. 4 are samples of the synthesized image patches by our model; we noted that the foreground nuclei are close to the original histopathology patches, and the background shows a realistic texture. Fig. 1 fourth row shows some examples of our synthesized patches, it shows that our model can produce clear nuclei boundary in the overlapped nuclei regions.

3.3. Application to nuclei segmentation

To demonstrate the effectiveness of the proposed Sharp-GAN, we evaluate segmentation models trained using synthetic images generated from Sharp-GAN. Two semantic segmentation networks, U-Net [18] and SegNet [19] are employed to segment the nuclei foreground from the background.

We compare the segmentation networks using different training sets. R contains only real images from the

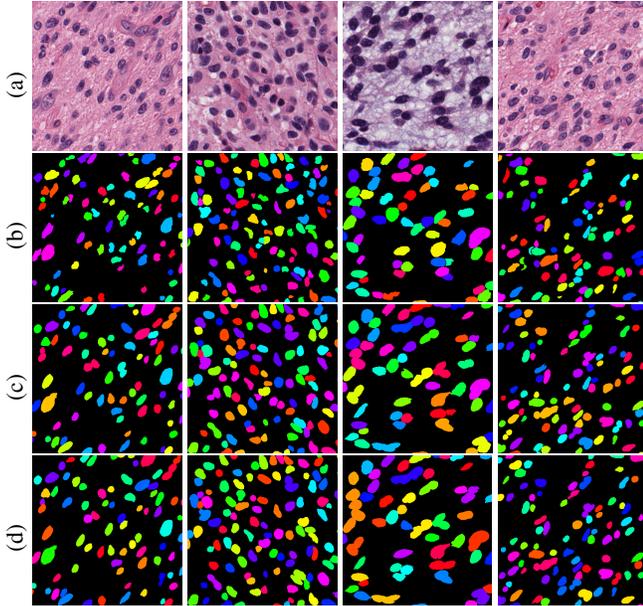


Fig. 5: Segmentation results. (a) Original images; (b) ground truth; (c) U-Net results using synthetic and real images; and (d) U-Net results using synthetic images only. Different colors represent different nuclei.

Table 2: Segmentation results on CPM-15&17 test set.

Model	Training	DQ	SQ	PQ	AJI
SegNet	R	0.765	0.759	0.584	0.588
	S (3k)	0.686	0.757	0.523	0.510
	R+S (3k)	0.780	0.769	0.602	0.608
	S (10k)	0.732	0.759	0.557	0.545
	R+S (10k)	0.806	0.772	0.608	0.623
U-Net	R	0.825	0.795	0.658	0.670
	S (3k)	0.755	0.738	0.562	0.593
	R+S (3k)	0.837	0.800	0.671	0.675
	S (10k)	0.804	0.788	0.634	0.616
	R+S (10k)	0.841	0.797	0.669	0.677

CPM15&17 training set; S denotes as set of pure synthetic images; and R+S denotes a combined set of real and synthetic images. As shown in Table 2, the results of both U-Net and SegNet trained using 3k synthetic images are only slightly worse than the results of models trained using real data. While we train the models using a combination of synthetic data and real data, their results of the four metrics are all higher than those of the models trained using only real images.

Meanwhile, we synthesize two datasets, one with 10k images and another with 3k images, and train U-Net and SegNet using them. Table 2 shows that the segmentation models trained using 10k synthetic images and real images outperform the models training using a combinations of 3k synthetic images and real images. Fig. 5 shows four segmentation examples of U-Net. They show that the model trained using only synthetic images is comparable to the model trained using both synthetic and real images.

4. CONCLUSION

In this paper, we propose the Sharp-GAN to generate realistic histopathology images. It uses nucleus distance map as input to clearly define the contours between overlapped and touching nuclei; and we propose the sharpness term to enhance the contrast of nuclei contours. Quantitative evaluation demonstrates that the proposed approach can generate histopathology images with high qualities. The segmentation results demonstrate that the models, e.g., U-Net and SegNet trained using both synthetic and real images significantly outperform the those trained using only real images.

5. REFERENCES

- [1] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter, “Segmentation of nuclei in histopathology images by deep regression of the distance map,” *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 448–459, 2018.
- [2] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot, “Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images,” *Medical Image Analysis*, vol. 58, pp. 101563, 2019.
- [3] Haotian Wang, Min Xian, and Aleksandar Vakanski, “Bending loss regularized network for nuclei segmentation in histopathology images,” in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2020, pp. 1–5.
- [4] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi, “A dataset and a technique for generalized nuclear segmentation for computational pathology,” *IEEE transactions on medical imaging*, vol. 36, no. 7, pp. 1550–1560, 2017.
- [5] Yanning Zhou, Omer Fahri Onder, Qi Dou, Efstratios Tsougenis, Hao Chen, and Pheng-Ann Heng, “Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation,” in *International Conference on Information Processing in Medical Imaging*. Springer, 2019, pp. 682–693.
- [6] Le Hou, Ayush Agarwal, Dimitris Samaras, Tahsin M Kurc, Rajarsi R Gupta, and Joel H Saltz, “Robust histopathology image analysis: To label or to synthesize?,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8533–8542.

- [7] Naiyun Zhou, Xiaxia Yu, Tianhao Zhao, Si Wen, Fusheng Wang, Wei Zhu, Tahsin Kurc, Allen Tannenbaum, Joel Saltz, and Yi Gao, "Evaluation of nucleus segmentation in digital pathology images through large scale image synthesis," in *Medical Imaging 2017: Digital Pathology*. International Society for Optics and Photonics, 2017, vol. 10140, p. 101400K.
- [8] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.
- [9] Augustus Odena, Christopher Olah, and Jonathon Shlens, "Conditional image synthesis with auxiliary classifier gans," in *International conference on machine learning*. PMLR, 2017, pp. 2642–2651.
- [10] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.
- [11] Aaron Hertzmann, Charles E Jacobs, Nuria Oliver, Brian Curless, and David H Salesin, "Image analogies," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, pp. 327–340.
- [12] Y.Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 2001, vol. 1, pp. 105–112 vol.1.
- [13] Quoc Dang Vu, Simon Graham, Tahsin Kurc, Minh Nguyen Nhat To, Muhammad Shaban, Talha Qaiser, Navid Alemi Koohbanani, Syed Ali Khurram, Jayashree Kalpathy-Cramer, Tianhao Zhao, et al., "Methods for segmentation and classification of digital microscopy tissue images," *Frontiers in bioengineering and biotechnology*, vol. 7, pp. 53, 2019.
- [14] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [15] Lin Zhang, Lei Zhang, Xuanqin Mou, and David Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.
- [16] Wufeng Xue, Lei Zhang, Xuanqin Mou, and Alan C Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–695, 2013.
- [17] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollár, "Panoptic segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9404–9413.
- [18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [19] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.