PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Generative modeling for renal microanatomy

Leema Krishna Murali, Brendon Lutnick, Brandon Ginley, John Tomaszewski, Pinaki Sarder

Leema Krishna Murali, Brendon Lutnick, Brandon Ginley, John E. Tomaszewski, Pinaki Sarder, "Generative modeling for renal microanatomy," Proc. SPIE 11320, Medical Imaging 2020: Digital Pathology, 113200F (16 March 2020); doi: 10.1117/12.2549891

SPIE.

Event: SPIE Medical Imaging, 2020, Houston, Texas, United States

Generative modeling for renal microanatomy

Leema Krishna Murali[1], Brendon Lutnick[2], Brandon Ginley[2], John E. Tomaszewski[2], and Pinaki Sarder[2,*]

[1]Department of Biomedical Engineering, SUNY Buffalo, Buffalo, NY, USA 14228,
[2]Department of Pathology and Anatomical Sciences, SUNY Buffalo, Buffalo, NY, USA 14203

[*]Address all correspondence to: Pinaki Sarder
Tel: 716-829-2265; E-mail: pinakisa@buffalo.edu

ABSTRACT

Generative adversarial networks (GANs) have received immense attention in the field of machine learning for their potential to learn high-dimensional and real data distribution. These methods do not rely on any assumptions about the data distribution of the input sample and can generate real-like samples from latent vector space based on unsupervised learning. In the medical field, particularly, in digital pathology expert annotation and availability of a large set of training data is costly and the study of manifestations of various diseases is based on visual examination of stained slides. In clinical practice, various staining information is required to improve the pathological diagnosis process. But when the sampled tissue to be examined is limited, the final diagnosis made by the pathologist is based on limited stain styles. These limitations can be overcome by studying the usability and reliability of generative models in the field of digital pathology. To understand the usability of the generative models, we synthesize in an unsupervised manner, high resolution renal microanatomical structures like renal glomerulus in thin tissue histology images using state-of-art architectures like Deep Convolutional Generative Adversarial Network (DCGAN) and Enhanced Super-Resolution Generative Adversarial Network (ESRGAN). Successful generation of such structures will lead to obtaining a large set of labeled data for further developing supervised algorithms for disease classification and understanding progression. Our study suggests while GAN is able to attain formalin fixed and paraffin embedded tissue image quality, GAN requires further prior knowledge as input to model intrinsic micro-anatomical details, such as capillary wall, urinary pole, nuclei placement, suggesting developing semi-supervised architectures by using these above details as prior information. Also, the generative models can be used to create an artificial effect of staining without physically tampering the histopathological slide. To demonstrate this, we use a CycleGAN network to transform Hematoxylin and eosin (H&E) stain to Periodic acid-Schiff (PAS) stain and Jones methenamine silver (JMS) stain to PAS stain. In this way GAN can be employed to translate different renal pathology stain styles when the relevant staining information is not available in the clinical settings.

Keywords: Digital pathology, Hematoxylin and eosin, Periodic acid-Schiff, Jones methenamine silver, Unsupervised learning, Generative Adversarial Network, Machine Learning, Glomeruli

1. INTRODUCTION

In standard histopathology practice, upto two chromogens can be applied on the same sample for brightfield microscopy [1]. Pathologists perform serial cuts of varying thickness of same tissue block and stain with different techniques and antibodies to extract multiple features [1]. This raises the issue of inter-slide variability [1] preventing cell-level segmentation across slides. As a consequence, different stained sections do not represent the same tissue sample. GANs gives an opportunity to obtain virtual staining of images with one dye from another from exactly the same sample, for example, PAS or Silver stained images from images dyed with widely available and less expensive H&E stain. This may lead to effective pathological examination and bypasses the procedure of restaining the tissue sample with other special stains for additional investigations, as the information obtained with different stains is complementary. Many important discoveries concerning diseases have been made pathologists and researchers by closely examining the pathological specimens. Also new clinicopathological relationships can be discovered [2]. This

Medical Imaging 2020: Digital Pathology, edited by John E. Tomaszewski,
Aaron D. Ward, Proc. of SPIE Vol. 11320, 113200F · © 2020 SPIE
CCC code: 1605-7422/20/$21 · doi: 10.1117/12.2549891

can be seen in the discovery of *H. pylori* by a pathologist who examined the gastric mucosa of patients with gastritis [2, 3].

In the clinical settings, human experts such as pathologists are primarily involved in the interpretation of medical images. However, due to large variations in pathology and potential fatigue of experts, computer assisted interventions have found their significance in the field of medical image analysis [2]. In the recent years, various attempts have been made to capture the stained specimen slide with a digital scanner and save it as whole slide image (WSI). In order to analyze large number of WSIs, machine learning (ML) algorithms have been employed to assist in image analysis and diagnosis. However, medical images like digital pathological images and tasks require some special processing techniques and specific characteristics that need to be considered. Some problems specific to pathological image analysis using machine learning algorithms include very large image size, different magnification levels, color variation and artifacts, etc. One of the compelling challenges that limit easy employment of ML methods in pathology is variability in the number of patterns derived from the tissues from a computational perspective [4]. This form of polymorphism requires many training cases of the inherent architecture of deep networks. Another key challenge is with regard to the variations between different patient types. There is no similarity between two samples for the same disease state as a result there is contradiction in which we notice that the same cell type has different characteristics in different patients [4]. To avoid bias that originates from the data used by the system to learn, we would need lot of training data in order to generalize the system to all patient types properly. In addition to this the deep networks have no established reason to explain why a specific decision was made when dealing with histopathological scans. This results in the lack of transparency in the path of diagnosis using ML algorithms. In this study, we focus on one of the most recent advances in deep learning, which is generative adversarial network (GAN) and its potential use cases of GANs in the medical domain.

GANs are a type of neural network model which is built of two networks namely generator and discriminator that are trained simultaneously. The generator is focused on generating realistic samples by learning the distribution of the training samples based on the feedback given by the discriminator. Adversarial training lies at the heart of the real conceptual progress for Machine Learning; particularly, for generative models. These models are estimated via an adversarial process to overcome the difficulty of approximating probabilistic computations [5]. The advantage of these models is that they can represent sharp and degenerate distributions of data over the generative stochastic networks that use Markov chains [5]. The latter requires a somewhat blurry distribution in order for the chains to mix between generator and discriminator models. One of the most successful implementations of generative adversarial network (GAN) is Deep Convolutional GAN (DCGAN) that have shown tremendous potential in recent years and have wide range of applications ranging from image synthesis to enhancing the image quality [6]. The DCGAN model's fundamental component is to replace the fully connected layers in the generator with upsampling convolutional layers. Convolutional networks in the generator and discriminator model help in finding deep correlation within an image, that is looking for spatial correlation thereby avoiding the need to meticulously select the network hyperparameters. This makes the network suitable to generate any medical image (in this study we are primarily focusing on histopathological images) reducing the investment on labelled data as required by other supervised machine learning algorithms. This form of synthetic data augmentation of high-resolution images enables more variability and enriches the dataset to improve the efficiency and accuracy of training algorithms [7]. This could expand the datasets that doctors and researchers have to work with when it comes to rare disease types, especially when the abnormal cases are very few which are needed to detect and diagnose. Moreover, GANs could assist to overcome privacy challenges that surround the use of patient data as the synthetic images are not associated to any patient and it's safer and more anonymous to transfer the data for any research purposes [8]. Additionally, with the help of recent advances in GANs [6], enhanced resolution images can be generated that may enhance the fine details on an image, thus making the network suitable for generating medical images. In this way GANs can alleviate the problem of data scarcity and class imbalance due to rare nature of some pathologists. In addition to image synthesis, most studies have employed GAN technique to create label-to-segmentation translation [9], segmentation-to-image translation [10], medical cross modality translations [11]. Although these networks achieve good levels of realism they pose few drawbacks that hinder their acceptance in the medical community. It has been studied that the typical GAN training is unstable resulting in convergence issues such as mode collapse [12]. The fundamental work of these models on computer vision datasets like facial datasets focuses on solving this problem but in medical imaging where the number of modes are unclear, identification of such uncertain and unstable situations can be very challenging.

1.1 Our Contribution

In this study, we explore the usability and reliability of generative models in the field of digital pathology. Our main goal will be not generating something realistic, but exploiting the inner knowledge of these generative models in pathology. Toward this goal, in this work, we have attempted to generate renal glomeruli images using GAN, and qualitatively studied the validity of the micro-anatomical structures in the generated images via expert renal pathologists. The generated large set of labeled data from a small set of labeled datasets can be potentially useful as an augmentation technique to prevent against supervised network overfitting. In addition to this we have also attempted to pursue unsupervised stain style translation on renal histopathological whole slide images for few special stains used in clinical settings. This process of using dedicated computer algorithms can generate other stained versions of a real slide stained using a traditional dyeing method to highlight relevant histological and cell features on the same slide.

2. METHODOLOGY

2.1 Computational Model

We have developed DCGAN combined with ESRGAN in Pytorch for synthesis of high-resolution renal histopathology images. The network design of DCGAN is similar to that of the Vanilla GAN after removing the fully connected layers and replacing all pooling layers with strided convolutions [6]. The non-linearity function used at every layer in the generator network and discriminator network is ReLU (Rectified Linear Unit) and Leaky-ReLU respectively except for the output layer [6]. The network was trained at a learning rate of 0.0002 and mini-batch size of 20 for 500 epochs. The input glomeruli images were resized to 512×512 pixels and fed into the DCGAN network.

In previous works [6], DCGAN was observed to be stable in generating images up to a resolution of 512×512. Therefore, in order to achieve higher resolution, ESRGAN was combined along with the DCGAN network thereby resulting in high resolution glomeruli images with size greater than 1024×1024 pixels. As a first step the network was trained to generate 64×64 resolution glomeruli images. Next to check for the stability of the network, it was trained with the same values of the hyperparameters (changing the expected reconstructed image size) to reconstruct 128×128 spatial resolution images. It was observed that the network attained its stability after around mid of 200 epochs with the generator loss dropping to zero and discriminator loss increasing to very high value.

To train the ESRGAN, the highest resolution images (512 × 512) generated by DCGAN were then fed to ESRGAN. This enhanced version of super resolution GAN adapts to relativistic average GAN instead of normal GAN, uses features' prior activation to improve the perceptual loss, incorporates a Residual-in-Residual Dense Block (RRDB) which is easier to train, and comprises of higher capacity than the normal residual block[13]. The generator network is built of 23 RRDB, each contains 3 RDB. The training process is divided into two stages. Namely, we train a peak signal-to-noise ratio (PSNR)-oriented model with 1-norm distance between recovered image and the ground-truth (L1 loss) for learning rate initialized to 2×10^{-4}, and this trained model is then employed as an initialization for the generator with learning rate 1×10^{-4}).

We have employed the state-of-art architecture CycleGAN [14] to achieve the unsupervised stain style transfer on patches of WSI. The behavior of the network is observed by inputting different stained WSI. The model is trained in an unsupervised way using collection of H&E (Silver) patches extracted from WSI as the source and PAS patches as the target domain. This model is built of two GANs, each has a discriminator and a generator model. Each GAN comprises of a conditional generator network that synthesize an image given an input image and discriminator model to predict how likely the generator image resembles to those from the target image collection [14]. This resemblance is quantified by cycle consistency loss [14] that represents the L1 norm between the input and the CycleGAN generated image. The model works by taking an input image from H&E collection which is fed to the first generator, transforming to an image in the target domain (PAS). This new generated image is then fed to another generator which converts back into an image from the original domain (H&E). The discriminator distinguishes between the original PAS patch and the generated PAS patch. This is way the discriminator is able to defy the generator and reject images generated by it. On the other hand, the generator tries to generate images close to the original images in the other domain by accepting the feedback given by the discriminator. In this way, Nash equilibrium is achieved when the distribution of the generated samples become close to the desired distribution.

The potential of these state-of-art deep networks to generate large-scale datasets at unprecedented levels of realism requires high computational power of specific accelerators. Training these deep networks using ordinary computers with central processing units is extremely time-consuming and impractical. Hence, it's obvious to use graphical processing units to deploy deep networks in practice in order to accomplish such complex tasks. For example, the team at NVIDIA presented a deep network called Progressive growing GAN (ProGAN) that was successful in generating detailed 1024 × 1024 images by training the network on 8 Tesla V100 GPUs for 4 days [15]. Pathology laboratories are already under immense financial pressure to adopt WSI technology that requires acquiring and storing gigapixel histopathological scans. In addition to this, investing on high power computational resources will be an additional burden. Additionally, expert knowledge is essential to fine tune the network parameters in order to achieve nearly stable training of such deep networks. It is essential to have mathematical understanding of the deep layers so that the network can be modified at the right layer in order to accomplish the desired goal.

2.2 Computational resource

All training for this work was conducted via the computational resources available at the center for computational research (CCR) at University at Buffalo. Hardware used for training the network is Dell PowerEdge R740 supported by two NVIDIA Volta Tesla V100 PCIe GPUs and 32x2.10GHz Intel Xeon Gold 6130 CPU. Total of 16 nodes (32 cores per node) were used with memory bandwidth 900 GB/sec.

3. DATA

The training set used DCGAN consisted of 20K PAS, H&E, and Trichrome stained glomeruli images (Figure 1) obtained from human biopsies of various pathology. The image staining, imaging, and data generation followed similar method as described in our previous publication [16].

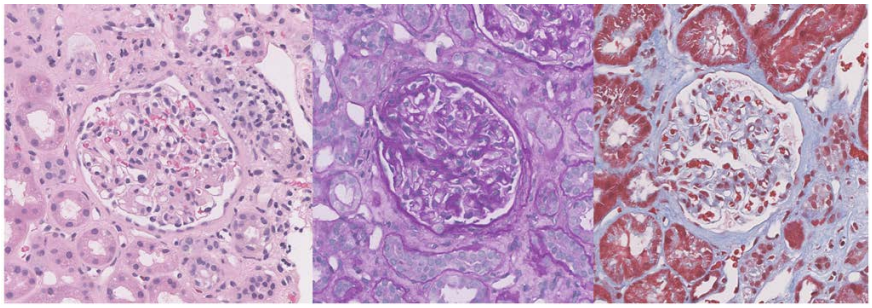

Figure 1. Training data: H&E (left), PAS (middle), and Trichrome (right) stained glomeruli images

Training data for CycleGAN: Patches of size 256×256 (Figure 2) are extracted from whole slide images (WSI) of H&E, PAS and Silver having a magnification of 40x. These patches are fed to the network for unpaired image-to-image translation from one stain style to another. A total of approximately 10,000 patches are extracted from one WSI that varies depending on the dimension of the each WSI.

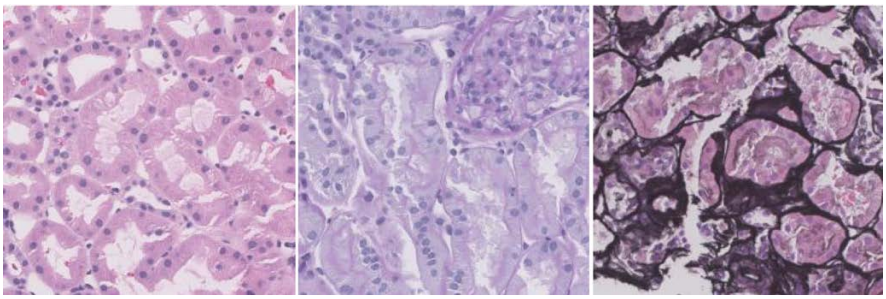

Figure 2. Training data: H&E (left), PAS (middle), and Silver stained patches from WSI

4. RESULTS

The results of our study are discussed in the following two sections:

4.1 Synthesis of high-resolution glomeruli images

The automatic segmentation of glomeruli from the resulting whole slide images was performed using H-AI-L segmentation tool developed in our previous work [17]. DCGAN is then trained with patches centered around the located glomeruli having spatial resolution of 1200 × 1200 pixels. The network was trained initially to generate 64 × 64 spatial resolution images (Figure 3), and later up to 512 × 512 resolution by adding additional convolutional layers. Training was successfully completed when the discriminator's loss was more than the generator, implying that the generator was successful in counterfeiting real images and deceiving the discriminator.

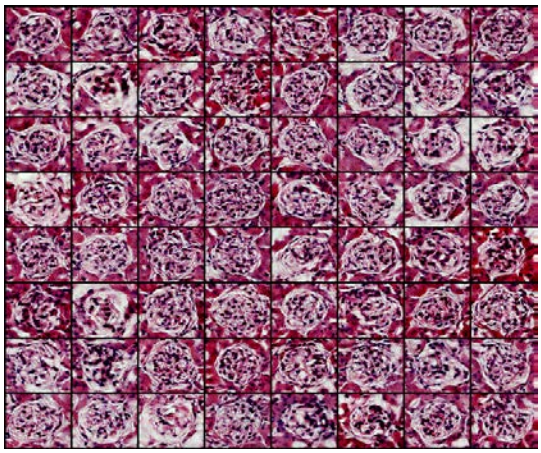

Figure 3. Generated glomeruli images of spatial resolution 64 × 64 obtained after training DCGAN for 448 epochs.

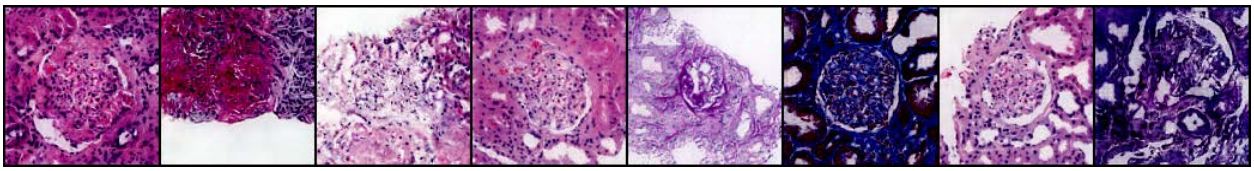

Figure 4. Generated glomeruli like images at spatial resolution 128 × 128. The network was trained for 2 days+ with size of the latent Z vector as 100.

The synthetic glomeruli images were provided to expert pathologists to study whether GAN is able to model the glomerular micro-anatomy. While the experts found that the reconstructed images achieve the similar quality as seen in formalin fixed paraffin embedded (FFPE) tissue sample images, they opined that the reconstructed images are unable to reconstruct fine details that have importance in renal disease pathology. Namely, the reconstructed images do not reconstruct well the capillary shape and wall, nuclei are often found to be smudged, some nuclei are visible as floating in the luminal space, as well as depict unusual shapes (Figure 5). Urinary pole is found to be incorrectly shaped and the epithelial cells are rendered inconsistently (Figure 6).

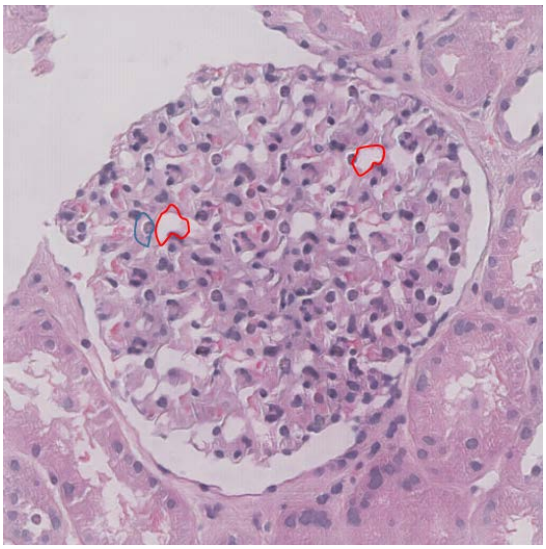

Figure 5. Generated glomerulus of spatial resolution 512 × 512. The network was trained with a total of 20,000 images for 3 days+ with the same size of the latent Z vector as discussed above. Highlighted red and blue regions in the above figure indicate unusual appearance of capillary wall and endothelial cells, respectively.

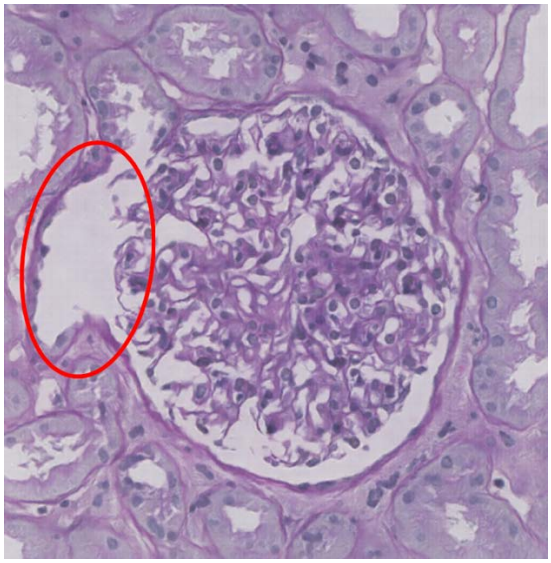

Figure 6. The red highlighted region in the above image indicates that the urinary pole (UP) is incorrectly shaped in the reconstructed glomerulus image. In addition, the capillaries in the Bowman's space depict unusual morphology.

The result of DCGAN was then fed to Enhanced Super Resolution GAN (ESRGAN) [18] to achieve glomeruli images in a magnified resolution (greater than 1024×1024) as shown in Figure 5. As before, while the image appears to be real-like, the experts opined that the fuschin quality in the reconstruction is poor. In addition, we estimate the Fréchet Inception Distance (FID) to evaluate the quality and diversity of the generated samples by GAN [19]. The activations of the last pooling layer from the Inception v3 model are used to summarize each image. In this way feature vectors are calculated for generated and real images. The FID score is then calculated using the equation [20].

$$d^2\big((m, C), (m_w, C_w)\big) = \| m - m_w \|_2^2 + Tr\left(C + C_w - 2(CC_w)^{1/2}\right) \quad (1)$$

where m and m_w refer to the feature-wise mean of the real and generated images. The covariance matrix for the real and generated images are C and C_w respectively. Tr refers to the trace linear algebra operation. The feature vectors are calculated for 10 real and synthetic images using the Inception v3 model. The FID score obtained using equation (1) is approximately 514 indicating higher variations between the synthetic and real data distributions.

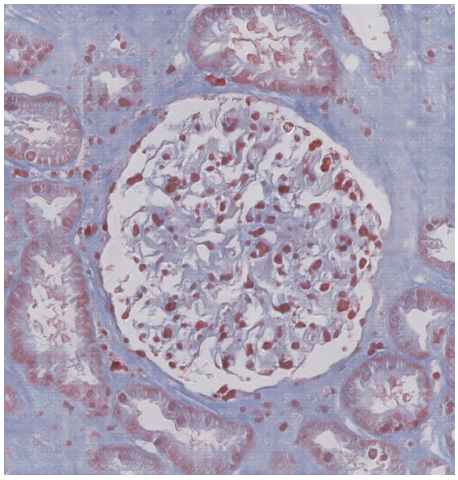

Figure 7. Generated glomerulus of spatial resolution 2024 × 2028.

4.2 Stain style transfer

The random patches are extracted from the resulting WSI using OpenSlide Python library. The down sample factor for each WSI is adjusted based on the objective power of individual WSI. The non-tissue patches are eliminated in this process of extraction by summing the pixel values in the mask generated for the WSI. Only those patches are retained for which the resulting sum of pixels in the mask equals to a non-zero value. The patches are initially normalized to a range between -1 and +1 to speed up the learning process leading to faster convergence. The model was trained with random H&E patches extracted from WSI for about 200 epochs. This learning process for each epoch continues until the training set gets exhausted.

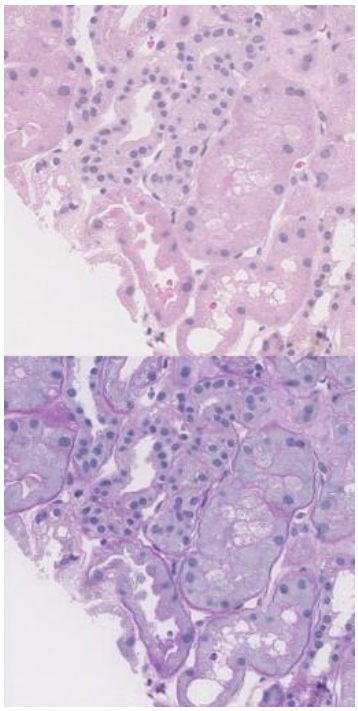

Figure 8. Original H&E patch (top) and generated PAS patch (bottom).

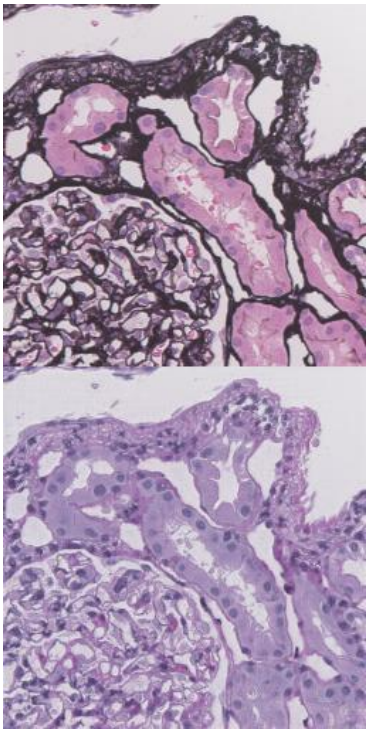

Figure 9. Original Silver patch (top) and generated PAS patch (bottom).

5. CONCLUSIONS

GAN is shown to be successful in producing glomeruli images achieving the quality as seen in FFPE tissue images given a set of unlabeled data. Focusing on the fine structures of the glomerulus, GANs are unable to reconstruct finer details that have importance in renal disease pathology. As a next step, the structures within the glomerulus that are significant with respect to the disease pathology would be fed to the network as an additional information prior to training. As a result, the network would learn these meticulous details, thereby improving the performance of reconstruction with respect to finer details seen in disease pathology. Although there is tremendous success of artificial intelligence (AI) in computer vision applications, there are clearly many challenges that need to be overcome in order to exploit deep learning techniques in digital pathology. The adoption of deep learning in pathology certainly requires expertise knowledge, support of computational processing power and big data repositories. Nonetheless, should these challenges be overcome, AI could prove to be an incredibly useful tool for improving the diagnosis of diseases and productivity of pathologists.

6. ACKNOWLEDGEMENTS

The project was supported by the faculty startup funds from the Jacobs School of Medicine and Biomedical Sciences, University at Buffalo; Buffalo Blue Sky grant, University at Buffalo; NIDDK Diabetic Complications Consortium grant DK076169; NIDDK grant R01 DK114485 & DK114485 02S1; NIDDK CKD Biomarker Consortium grant U01 DK103225; and NIDDK Kidney Precision Medicine Project grant U2C DK114886. We acknowledge the assistance of the Multispectral Imaging Suite and Histology Core Laboratory in the Dept. of Pathology & Anatomical Sciences, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo.

REFERENCES

1. Zhaoyang Xu, C.F.M., Béla Bozóky, Qianni Zhang, *GAN-based Virtual Re-Staining: A Promising Solution for Whole Slide Image Analysis.* arXiv:1901.04059, 2019.
2. Daisuke Komura, S.I., *Machine Learning Methods for Histopathological Image Analysis.* Computational and Structural Biotechnology Journal, 2018. **16**: p. 34-42.
3. Marshall, B., *A brief history of the discovery of Helicobacter pylori.* Springer, Tokyo, 2016: p. 3-15.
4. Pantanowitz, H.R.T.a.L., *Artificial Intelligence and Digital Pathology: Challenges and Opportunities.* Journal of Pathology Informatics, 2018.
5. Ian J. Goodfellow, J.P.-A., Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. *Generative Adversarial Networks*. in *NIPS*. 2014.
6. Alec Radford, L.M., Soumith Chintala, *Unsupervised Representation Learning With Deep Convolutional Generative Adversarial Networks*, in *International Conference on Learning Representations 2016*. 2016, January 7: Caribe Hilton, San Juan, Puerto Rico.
7. Maayan Frid-Adar, I.D., Eyal Klang, Michal Amitai, Jacob Goldberger, Hayit Greenspan, *GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification.* arXiv:1803.01229, 2018.
8. Ridley, E.L., *AI algorithm produces synthetic brain MR images*. 2018, AuntMinnieEurope.
9. P. Costa, A.G., M. I. Meyer, M. D. Abramoff, M. Niemeijer, ` A. M. Mendonc¸a, and A. Campilho, *Towards adversarial retinal image synthesis.* arXiv preprint arXiv:1701.08974, 2017.
10. W. Dai, J.D., X. Liang, H. Zhang, N. Dong, Y. Li, and E. P. Xing, *Scan: Structure correcting adversarial network for chest x-rays organ segmentation.* arXiv preprint arXiv:1703.08770, 2017.
11. D. Nie, R.T., C. Petitjean, S. Ruan, and D. Shen, *Medical image synthesis with context-aware generative adversarial networks.* arXiv preprint arXiv:1612.05362, 2016.
12. Antonia Creswell, T.W., Vincent Dumoulin, Kai Arulkumaran, Biswa and a.A.A.B. Sengupta, *Generative adversarial networks: An overview.* IEEE Signal Processing Magazine, 2018. **35(1)**: p. 53-65.

13. Wang, X., et al., *ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks.* eprint arXiv:1809.00219, 2018: p. arXiv:1809.00219.
14. Jun-Yan Zhu, T.P., Phillip Isola, Alexei A. Efros, *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.* arXiv:1703.10593v6, 2018.
15. Tero Karras, T.A., Samuli Laine, Jaakko Lehtinen, *Progressive Growing of GANs for Improved Quality, Stability, and Variation.* arXiv:1710.10196, 2018.
16. Ginley, B., et al., *Computational Segmentation and Classification of Diabetic Glomerulosclerosis.* Journal of the American Society of Nephrology, 2019: p. ASN.2018121259.
17. Lutnick, B., Ginley, B., Govind, D. et al., *An integrated iterative annotation technique for easing neural network training in medical image analysis.* Nature Machine Intelligence, (2019)(1(2)): p. 112-119.
18. Xintao Wang, K.Y., Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, Xiaoou Tang, *ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks*. 2018, September 17, Cornell University.
19. Martin Heusel, H.R., Thomas Unterthiner, Bernhard Nessler, Sepp Hochreiter, *GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium.* arXiv:1706.08500, 2017.
20. Landau., D.C.D.a.B.V., *The Fréchet distance between multivariate normal distributions.* Journal of Multivariate Analysis, 1982. **12**: p. 450-455.