

# Histopathology DatasetGAN: Synthesizing Large-Resolution Histopathology Datasets

*S. Rizvi<sup>1</sup>, P. Cicalese<sup>1</sup>, S. Seshan<sup>3</sup>, S. Sciascia<sup>4</sup>, J. Becker<sup>2</sup> and H. Nguyen<sup>1</sup>*

1. Department of Electrical and Computer Engineering, University of Houston, Houston, TX
2. Institute of Pathology, University Hospital of Cologne, Cologne, Germany
3. Weill Cornell Medicine, Cornell University, New York City, New York
4. Center of Research of Immunopathology and Rare Diseases, University of Turin, Turin, Italy  
{srizvi10, pcicalese, hienvnguyen}@uh.edu, sv2002@med.cornell.edu  
savino.sciascia@unito.it, janbecker@gmx.com

## I. INTRODUCTION

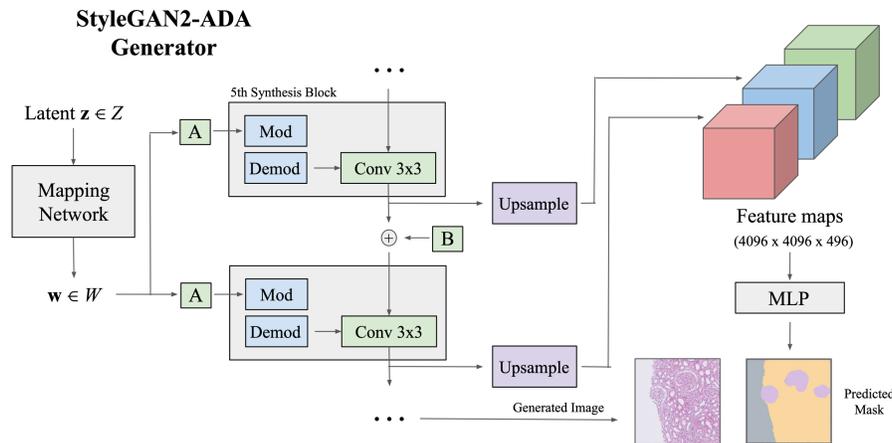
Deep learning-based methods have powered recent advancements in medical image segmentation, accelerating the field past previous statistical and Machine Learning-based methods [1]. This, however, has simultaneously created a need for large quantities of labeled data, which is difficult in domains such as medical imaging where labeling is expensive and requires expert knowledge. Semi-supervised learning (SSL) addresses these limitations by augmenting labeled data with large quantities of more widely available unlabeled data. Existing semi-supervised frameworks based on pseudo-labeling [2] or contrastive methods [3], however, struggle to scale to the high resolution of medical image datasets. In this work, we propose the Histopathology DatasetGAN (HDGAN) framework, an extension of the DatasetGAN framework for image generation and segmentation that scales well to large-resolution histopathology images. We make several adaptations on the original framework, including updating the generative backbone, selectively extracting latent features from the generator, and switching to memory-mapped arrays. These changes reduce the memory consumption of the framework, improving its applicability to medical imaging domains.

Previous works have shown that generative models learn a powerful, detailed latent space representation while learning to generate images [4]. DatasetGAN [4] leverages this property for semantic segmentation by performing pixel-wise classification on the latent features of Generative Adversarial Networks (GANs) to generate segmentation maps. Our work focuses on improving upon the DatasetGAN framework to increase its applicability and computational feasibility on high-resolution medical image datasets.

## II. METHODS

The dataset for our experiments consisted of 1,577 whole slide images (WSIs) taken from 100 native kidney biopsies; 50 with a diagnosis of thrombotic microangiopathy (TMA) and 50 mimickers of TMA, with similar or overlapping histopathological features in which a different diagnosis was made. The WSIs are taken from the three medical centers (Cologne, Weill-Cornell and Turin) and included at least three of the four diverse histopathological stainings: hematoxylin-eosin (HE), periodic-acid Schiff (PAS), Jones silver and trichrome, all scanned with a x40 objective. In this work, we pre-process the WSIs by first extracting 4096x4096 tiles which contained morphological compartments of interest. We then filtered out tiles which consisted of mostly whitespace, resulting in a dataset of 32,732 tiles. We split the tiles into 5 folds and trained our StyleGAN model on 4 of the folds, leaving one out as a holdout.

Our framework builds upon the original DatasetGAN framework proposed in [4]. The core idea is, given a latent vector  $z$  that is passed through the generator, we can extract the latent feature activations from convolutional layers of the generative network. These feature maps contain semantic information about the image being generated, and therefore are useful for pixel-wise segmentation tasks. The feature maps, which

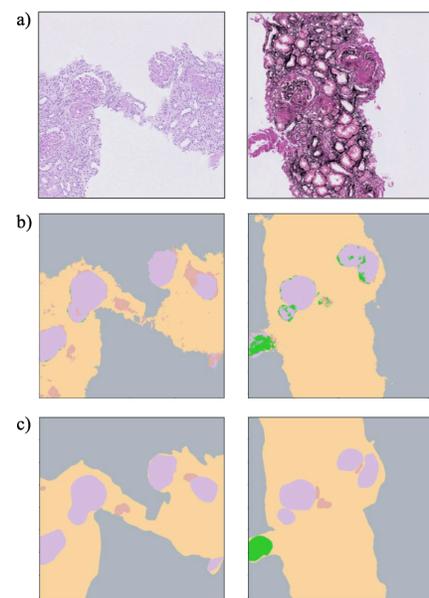


**Figure 1.** Histopathology DatasetGAN framework overview.

are extracted at different resolutions from the generator network, are upsampled to match the output image size and concatenated, yielding a single three-dimensional feature map tensor. With annotations on a small number of generated images, a small classifier is then trained to predict segmentation maps in a pixel-wise fashion from the extracted feature maps.

We extend this framework to high-resolution histopathology images by making several adjustments. First, we use StyleGAN2-ADA [5] as our generative backbone, which has proven to work well in limited data settings and at higher resolutions. Second, unlike [4], we choose to only extract and upsample latent features from the last 5 blocks in the synthesis module of the generator network, as shown in Figure 1, yielding feature maps which contain more fine-grained semantic information. Utilizing only these later features is sufficient for our pixel-wise segmentation task, and drastically reduces the amount of computation and upsampling required in the framework. To further decrease memory consumption during training on large histopathology images, we use memory-mapped arrays to load only portions of the feature map tensor currently in use. This allows us to scale our framework to 4096 x 4096 TMA tile segmentation, with samples shown in Figure 2. For our pixel-wise classifier, we use a 3-layer Multilayer Perceptron (MLP) with ReLU activations, Batch Normalization [6], and Dropout [7]. We train using vanilla stochastic gradient descent with a learning rate of 0.0001.

We evaluate the HDGAN framework on a 5-class semantic segmentation task on the TMA tile dataset. First, we generate 500 images with a truncation of 0.7 using our trained StyleGAN2-ADA network, saving the corresponding latent vectors. 36 of the 500 images were then selected for our dataset by an expert nephropathologist, who provided pixel-wise annotations of each morphological compartment for these 36 images. We randomly chose 16 images as the training set for the pixel-level classifiers, leaving 4 images for validation and the remaining 16 images as the test set. For baseline experiments, we refer readers to the original DatasetGAN paper [4] for comparisons to semi-supervised baselines.



**Figure 2.** Examples of synthesized images from HDGAN framework. (a) shows generated tiles from StyleGAN, (b) predicted segmentation map, and (c) ground truth segmentation.

### III. RESULTS

In Table 1 we report the class-wise pixel-level accuracy of the pixel-classifier on our test set. As expected, the pixel-level classifiers achieve high classification accuracy on several morphological compartments, with an average dice coefficient of 0.92 on the test set images. The classifiers struggle, however, when morphological structures in generated images resemble two different compartments, often misclassifying arteriole pixels as either glomeruli or artery. This semantic blending was confirmed by our expert annotator, sometimes forcing the annotator to choose between different class labels when a generated compartment had features of both classes.

**Table 1.** Class Pixel-wise Accuracy

| Class                       | Accuracy (%) |
|-----------------------------|--------------|
| Whitespace                  | 68.42%       |
| Cortical Tubulointerstitium | 66.57%       |
| Glomerulus                  | 82.89%       |
| Arteriole                   | 52.09%       |
| Artery                      | 77.82%       |

### IV. FUTURE WORK

The HDGAN framework suffers limitations when the generative backbone is not trained well on the original dataset. This manifested itself in the produced segmentation maps as noisy contours around compartments, where the semantic meaning held in the feature space of the generative networks was not decisively one compartment or another. Future work exploring the latent space of generative networks in relation to semantic meaning of morphological compartments could help disentangle the feature space and result in cleaner segmentation maps. Alternatively, future work could explore directly producing the segmentation map as an output of the generative model.

### V. ACKNOWLEDGMENTS

The research in this work was supported by the National Science Foundation and the National Institute of Health. The content represents the views of the authors, not necessarily the views of the funding agencies.

### VI. REFERENCES

- [1] T. Lei, R. Wang, Y. Wan, X. Du, H. Meng, and A. Nandi, "Medical image segmentation using deep learning: A survey." 2020.
- [2] D. Berthelot, N. Carlini, I. Goodfellow, N. Papernot, A. Oliver, and C. Raffel, "Mixmatch: A holistic approach to semi-supervised learning," *Advances in neural information processing systems*, vol. 32, 2019.
- [3] A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," *Advances in neural information processing systems*, vol. 30, 2017.
- [4] Y. Zhang, H. Ling, J. Gao, K. Yin, J. Lafleche, A. Barriuso, A. Torralba, and S. Fidler, "Datasetgan: Efficient labeled data factory with minimal human effort," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 10 145–10 155.
- [5] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12 104–12 114, 2020.
- [6] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *International conference on machine learning*. PMLR, 2015, pp. 448–456.
- [7] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

# Histopathology DatasetGAN: Synthesizing Large-Resolution Histopathology Datasets

S. A. Rizvi<sup>1</sup>, P. Cicalese<sup>1</sup>, S. V. Seshan<sup>3</sup>, S. Sciascia<sup>4</sup>, J. U. Becker<sup>2</sup>, H. V. Nguyen<sup>1</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, University of Houston, Houston, TX

<sup>2</sup>Institute of Pathology, University Hospital of Cologne, Cologne, Germany

<sup>3</sup>Weill Cornell Medicine, Cornell University, New York City, New York

<sup>4</sup>Center of Research of Immunopathology and Rare Diseases, University of Turin, Turin, Italy

## INTRODUCTION

- Deep Learning-based image generation and segmentation networks hold great potential for augmenting and segmenting scarce histopathology image datasets
- One of the most pressing limits is the need for large quantities of labeled histopathology data
- Previous frameworks are too computationally expensive for high resolution histopathology images
- The aim of this work is to propose the semi-supervised Histopathology DatasetGAN framework, with adaptations for application to high-resolution images and scarce medical image datasets

## METHODS

- Our dataset consists of 1,577 whole slide images (WSIs) from 100 kidney biopsies, 50 with a diagnosis of thrombotic microangiopathy (TMA) and 50 mimickers of TMA
- We extract 4096x4096 resolution tiles from the WSIs, and remove tiles composed primarily of whitespace, yielding 32,732 tiles for training the generative network
- For semantic segmentation, we annotate 36 generated images with 5 morphological class labels. We use 16 of the images to train the pixel classifier, and hold out 4 and 16 images for validation and test, respectively
- We utilize the StyleGAN2-ADA generative network, which has proven to work well in limited data scenarios
- We pass a latent vector,  $z$ , through the StyleGAN2 network and extract feature maps after each convolutional layer in the last 5 synthesis blocks
- Feature maps are upsampled to match the original image resolution, and a 3-layer multilayer perceptron is applied to classify pixels into a segmentation mask
- HDGAN framework is illustrated in Figure 1

## ARCHITECTURE

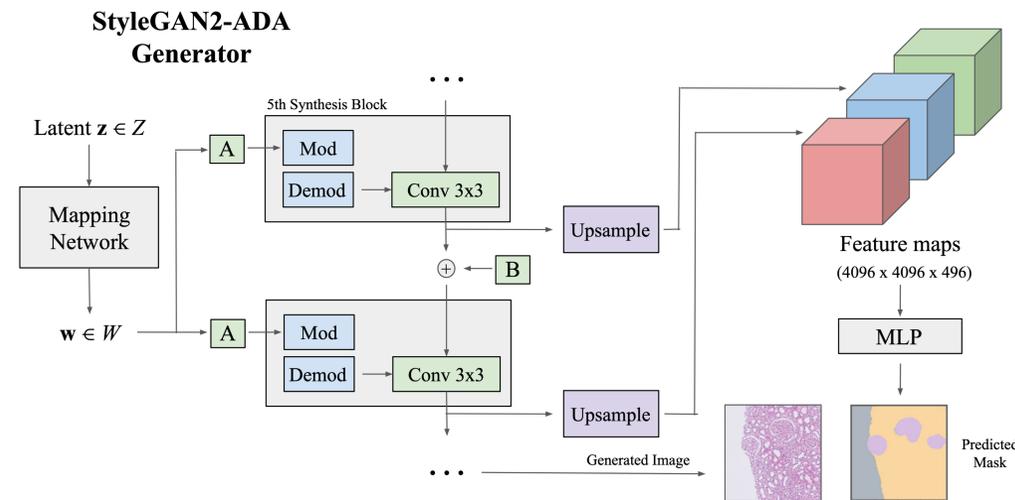


Figure 1. Histopathology DatasetGAN framework overview

## RESULTS

- Pixel-wise classifier achieves high accuracy on several morphological components in semantic segmentation task
- Average dice coefficient of 0.92 on test set images
- Predicted segmentation maps closely align with ground truth mask, with some semantic blending between morphological components (e.g., Glomeruli and Arteriole)

| Class                       | Accuracy (%) |
|-----------------------------|--------------|
| Whitespace                  | 68.42%       |
| Cortical Tubulointerstitium | 66.57%       |
| Glomerulus                  | 82.89%       |
| Arteriole                   | 52.09%       |
| Artery                      | 77.82%       |

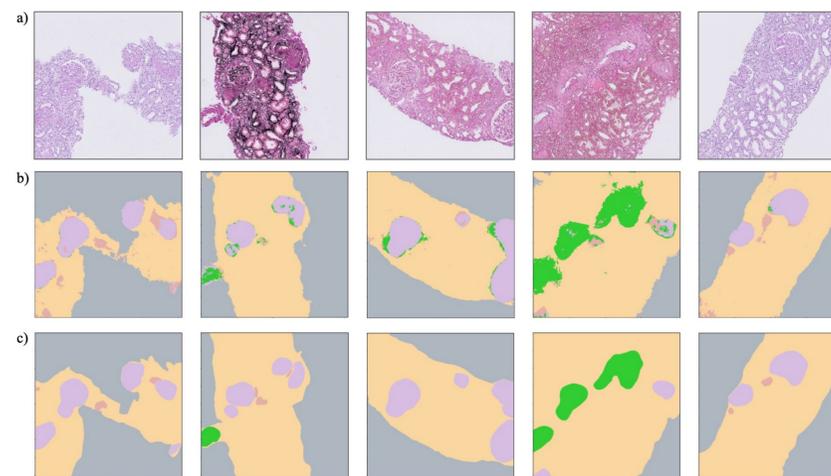


Figure 2. Examples of synthesized images (a) from HDGAN framework, alongside predicted segmentation maps (b) and ground truth masks (c).

## CONCLUSION

- Using a subset of latent features, the HDGAN framework successfully predicts segmentation masks for kidney biopsy images
- Glomeruli and arteries had the highest pixel-wise segmentation accuracy

## FUTURE WORK

- Explore latent space of generative network to investigate noisy contours in predicted masks, where semantic meaning in the feature space does not decisively indicate one class or another
- Apply methodology to other medical image datasets
- Explore developing a generative architecture which directly produces segmentation maps alongside images as output

## REFERENCES

- T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12 104–12 114, 2020.
- Y. Zhang, H. Ling, J. Gao, K. Yin, J. Lafleche, A. Barriuso, A. Torralba, and S. Fidler, "Datasetgan: Efficient labeled data factory with minimal human effort," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 10 145–10 155.

## ACKNOWLEDGEMENTS

The research in this work was supported by the National Science Foundation and the National Institute of Health. The content represents the views of the authors, not necessarily the views of the funding agencies