

# Diagnosability of Synthetic Retinal Fundus Images for Plus Disease Detection in Retinopathy of Prematurity

Aaron S. Coyner BS<sup>1</sup>, Jimmy Chen BA<sup>2</sup>, J. Peter Campbell MD<sup>2</sup>, Susan Ostmo MS<sup>2</sup>,  
Praveer Singh PhD<sup>3,4</sup>, Jayashree Kalpathy-Cramer PhD<sup>3,4</sup>, Michael F. Chiang MD, MA<sup>1,2</sup>

<sup>1</sup>Medical Informatics & Clinical Epidemiology, and <sup>2</sup>Ophthalmology  
Oregon Health & Science University, Portland, OR, United States;

<sup>3</sup>Radiology, MGH/Harvard Medical School, Charlestown, MA, United States;

<sup>4</sup>MGH & BWH Center for Clinical Data Science, Boston, MA, United States

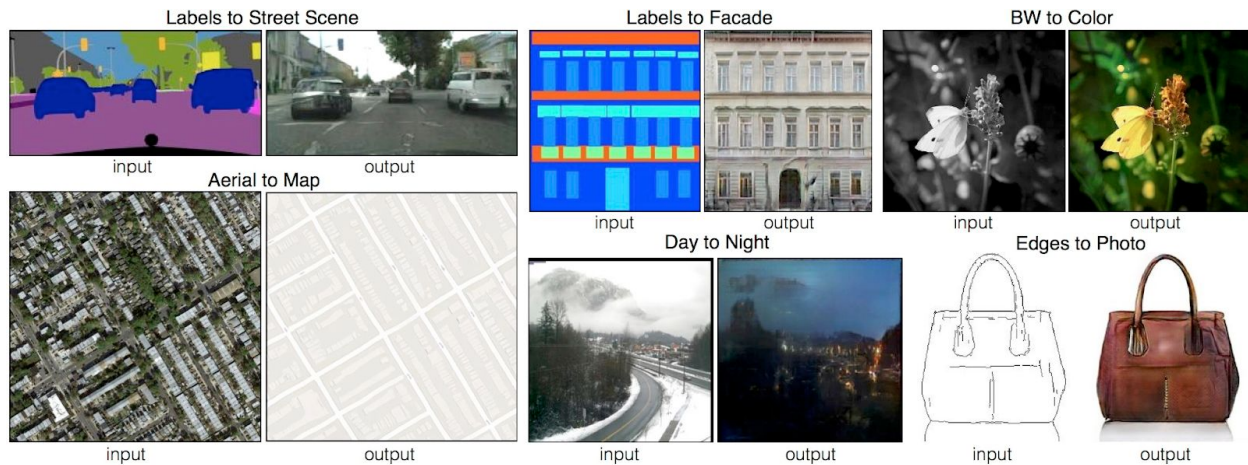
## Abstract

*Advances in generative adversarial networks have allowed for engineering of highly-realistic images. Many studies have applied these techniques to medical images. However, evaluation of generated medical images often relies upon image quality and reconstruction metrics, and subjective evaluation by laypersons. This is acceptable for generation of images depicting everyday objects, but not for medical images, where there may be subtle features experts rely upon for diagnosis. We implemented the pix2pix generative adversarial network for retinal fundus image generation, and evaluated the ability of experts to identify generated images as such and to form accurate diagnoses of plus disease in retinopathy of prematurity. We found that, while experts could discern between real and generated images, the diagnoses between image sets were similar. By directly evaluating and confirming physicians' abilities to diagnose generated retinal fundus images, this work supports conclusions that generated images may be viable for dataset augmentation and physician training.*

## Introduction

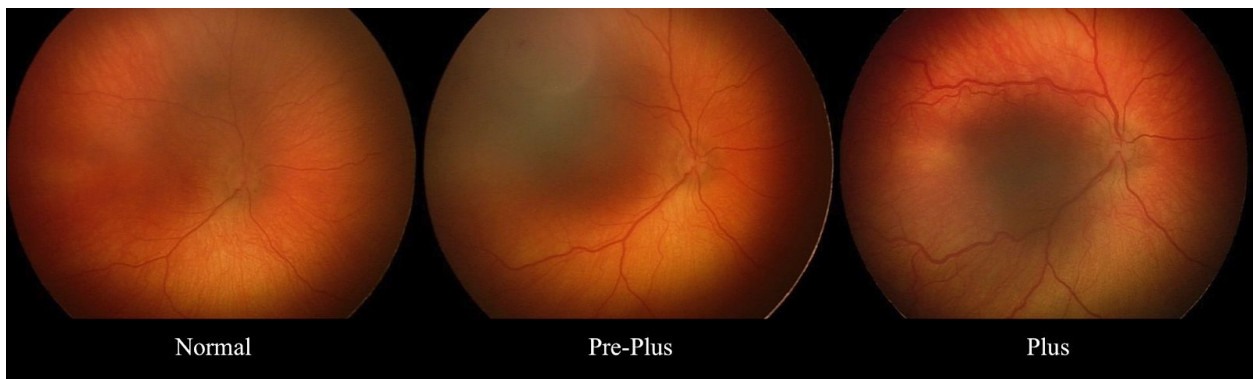
Advances in graphics processing units have allowed for development of complex models, such as deep neural networks and variants thereof.<sup>1,2</sup> Generative adversarial networks (GAN) are one such variant. These models contain both discriminative and generative networks that are trained to deceive one another.<sup>3,4</sup> A discriminative network attempts to estimate an output,  $y$ , given a set of inputs,  $x$ .<sup>3</sup> In contrast, a generative network attempts to model the distribution of  $x$  given  $y$ . To train these networks, data are supplied in pairs – inputs and their corresponding output(s). These two models are pitted against one another, and as training progresses, the ability of each model improves (i.e., as the discriminator better learns to discern between real and generated images, the generator must also learn how to better simulate generated data). Ideally, this results in a generator that consistently fools a well-trained discriminator into classifying its outputs as real. These models can be used for many types of data, but are primarily used for image generation. For images, both the discriminative and generative networks attempt to learn the overall style and pattern of output images (i.e., the relevant features of output images). However, the generative network also tries to learn how to map the original input image to the style of the output image. These models have begun to gain traction for synthesis of medical images.<sup>5-8</sup>

The model presented in Image-to-Image Translation with Conditional Adversarial Networks, “pix2pix,” allows one to employ style transfer (**Figure 1**).<sup>4</sup> This method has been used to map real images to labels, labels to images, convert black-and-white images to color, and convert images captured during the day to representations of the same images at night. A few studies have even used this method to map retinal blood vessel maps to retinal fundus images for research in diabetic retinopathy.<sup>8-10</sup> However, while these generated images were evaluated using subjective visual quality inspections and various image quality/reconstruction metrics, the diagnosability of said images – arguably the most important factor – was never formally evaluated.



**Figure 1: Example implementations of the pix2pix generative adversarial network.** This model demonstrates excellent ability to convert feature maps to real images (Labels to Street Scene, Labels to Facade, Edges to Photo), and real images to feature maps (Aerial to Map). The results are realistic and of relatively-high resolution. Figure adapted from *Image-to-Image Translation with Conditional Adversarial Networks*.<sup>4</sup>

In this study, we have deployed the pix2pix pipeline for retinal fundus image synthesis from retinal vessel maps.<sup>4</sup> The application of this work is in retinopathy of prematurity (ROP), a potentially-blinding disorder that affects premature infants.<sup>11,12</sup> A significant predictor of treatment-requiring ROP is the presence of plus disease, described as venous dilation and arterial tortuosity (**Figure 2**).<sup>12</sup> It stands to reason that, according to the definition of plus disease, the only information required to diagnose plus disease is the appearance of the major retinal blood vessels.<sup>12</sup> Herein, we describe two pix2pix models: one that generates realistic retinal fundus images, and another that only retains the major retinal blood vessels. To accomplish this task, we first generate retinal vessel maps from retinal fundus images using a previously-reported U-Net model, then create new retinal fundus images of varying pigmentations from original images using the raw retinal vessel maps, and also create new retinal fundus images of varying pigmentations that lack choroidal blood vessel patterns and/or any unique/abnormal features of the retina (e.g., haemorrhages, discolorations, etc.) using filtered retinal vessel maps.<sup>13,14</sup> Data for this study were obtained through the multi-center, NIH-funded, Imaging and Informatics in ROP (i-ROP) study centered at Oregon Health & Science University (OHSU).



**Figure 2: Example retinal fundus images.** From left to right, retinal fundus images of an eye that was originally diagnosed normal, developed pre-plus disease, and then plus disease. In plus disease images, retinal blood vessels are dilated and tortuous as compared to normal images. The degree of dilation and tortuosity of pre-plus blood vessels is less than that of plus disease blood vessels, but greater than normal.

## Methods

### *Institutional Review Board*

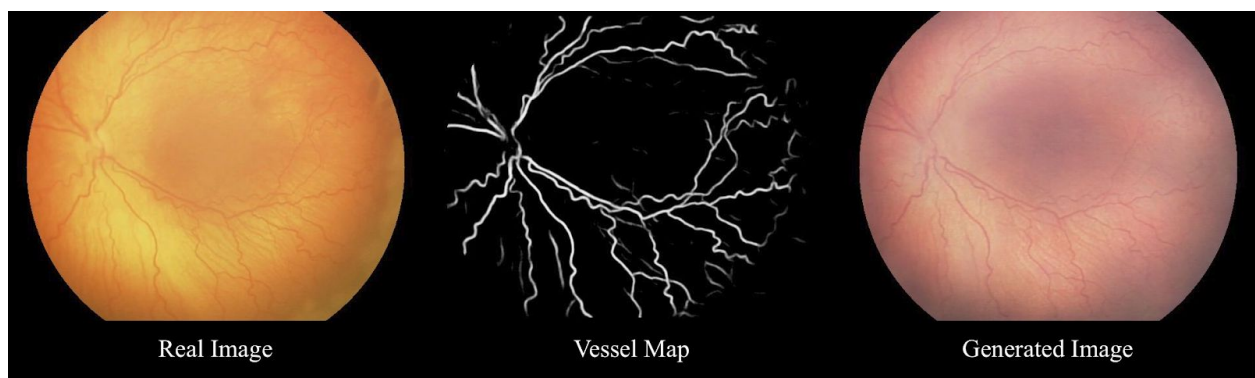
This study was approved by the Institutional Review Board at the coordinating center (OHSU) and at each of 8 study centers (Columbia University, University of Illinois at Chicago, William Beaumont Hospital, Children's Hospital Los Angeles, Cedars-Sinai Medical Center, University of Miami, Weill Cornell Medical Center, Asociación para Evitar la Ceguera en México [APEC]). This study was conducted in accordance with the Declaration of Helsinki. Written informed consent for the study was obtained from parents of all infants enrolled.

### *Retinal Fundus Image Dataset*

As part of the multicenter ROP cohort study, i-ROP, over 30,000 nasal, temporal, inferior, superior, and posterior-pole retinal fundus images were collected from preterm infants during routine ROP screening examinations. Between three and eight independent experts labeled each image set as normal, pre-plus, or plus, and an expert consensus diagnosis was formed and established as the ground truth diagnosis. Experts were all ophthalmologists with extensive experience in both ophthalmoscopic and image-based diagnosis of ROP. A subset of fundus images were selected; exclusion criteria were: images not centered on the posterior pole, images of stage 4 ROP (partial retinal detachment), and images of stage 5 ROP (total retinal detachment). The remaining images were downsampled to create an even distribution among babies diagnosed as normal, pre-plus, and plus, to a final dataset consisting of 594 wide-angle fundus images centered on the posterior pole. This dataset was randomly split (retaining the even distribution), 80/10/10, into train, validation, and test datasets, respectively. Because a subject may be represented in the dataset more than once (multiple imaging sessions), it was ensured that subjects were unique to each dataset.

### *Model Setup and Training*

Models were built and trained in Python using PyTorch on an Nvidia V100 GPU (Santa Clara, CA).<sup>15</sup> For each image in the training, validation, and test datasets, vessel maps were generated using a previously-trained U-Net.<sup>13,14</sup> The opensource pix2pix code was forked from a Github repository hosted by its authors.<sup>16</sup> We applied a modified pix2pix GAN, using ResNet9 blocks, to the i-ROP training dataset (**Figure 3**). The value of  $\lambda$  was set at 10; this weights the L1 loss 10 times greater than the adversarial loss of the generator during training, resulting in objectively higher fidelity images. Additionally, the original model was designed to generate color images of size 256 x 256, but retinal fundus images are generally color images of size 640 x 480. Rather than upsampling generated 256 x 256 images, which resulted in slightly blurred images that could (A) affect diagnosability and/or (B) be more easily discerned as a generated images, the model was modified to produce images of size 640 x 480. During training, each image in an image set (vessel map and corresponding retinal image) was scaled to size 536 x 536, and a random 480 x 480 crop was acquired from the same location for each image in the set. When images were generated, the 480 x 480 output image was resized to 640 x 480 to match the size of retinal fundus images. Finally, a black, circular mask was applied around the outside of the image to better mimic the appearance of images captured by retinal fundus cameras.



**Figure 3: Retinal image generation process.** Blood vessels of real images (left) are segmented and converted into retinal vessel maps (center). Pix2pix is used on raw or filtered retinal vessel maps (no discernible difference in image appearance) to generate images with similar vascular patterns (right).

Two separate models were trained: one on raw, grayscale vessel maps produced by the U-Net model (pix2pix-raw), and the other on the same vessel maps thresholded at pixel values greater than 25 (pix2pix-filtered). After the first training session, it was noted that although the U-Net was only trained to segment major retinal blood vessels, the reconstructed images contained similar choroidal blood vessel patterns. Upon further investigation, it was found that choroidal blood vessels were segmented, but at pixel intensities indistinguishable from the background to the human eye (pixel intensity < 26). Therefore, for the second training iteration, pixel values less than or equal to 25 were set to 0 to remove information about choroidal blood vessel patterns. The models were trained for 1,000 epochs using the Adam optimizer with a  $\beta$  value of 0.5. During the first 500 epochs, the learning rate was constant at  $2 \times 10^{-4}$ , and was linearly decayed to 0 over the remaining 500 epochs. Discriminator and generator loss functions on both the training and validation test sets were monitored to ensure learning was occurring at an equal rate between objective functions, and that overfitting was not occurring. The quality of image reconstructions were evaluated using the structural similarity index (SSIM) and the peak signal to noise ratio (PSNR), metrics often used to describe the quality of image reconstruction.<sup>17,18</sup>

### *Image Grading*

Of the 36 true retinal fundus images in the test dataset, 30 images were randomly selected for grading. Raw vessel maps and thresholded vessel maps were generated for each real image and used to generate reconstructions from their respective models. In total, there were three image sets (90 images) to be graded: 30 ground truth retinal fundus images, 30 reconstructions from pix2pix-raw, and 30 reconstructions from pix2pix-filtered. Using a custom, online grading system, a set of 3 independent ROP experts graded each image as normal, pre-plus, or plus, and also assessed whether they believed the image was real or generated.<sup>19</sup> All images were presented at a resolution of 640 x 480 x 3. In the event of a three-way tie for normal, pre-plus, or plus, the image was classified as pre-plus. The majority diagnoses of real images were used to compare agreement of diagnoses to generated image sets.

### *Data Analysis*

All analyses were performed in R. Majority diagnoses were determined for all images in a set, in addition to a majority vote on whether images were real or generated. Pearson's Chi-squared Test of Independence ( $\chi^2$ ) was used to determine if experts were statistically able to identify generated images from real images. Because some cell values of the confusion matrix for the Chi-squared test were less than five, a Fisher's Exact Test for Count Data was used to confirm findings. In order to determine if expert diagnoses were affected by generated images, the Cochran-Mantel-Haenszel test was used to compare the pix2pix-raw and pix2pix-filtered contingency tables to the real image contingency table. Further, Cohen's kappa ( $\kappa$ ) was used to measure agreement of diagnoses between generated images and real images.

## **Results**

The [discriminator, generator] losses on the training dataset for pix2pix-raw and pix2pix-filtered were [0.315, 0.224] and [0.170, 0.078], respectively. The [train, validation] PSNR values for pix2pix-raw and pix2pix-filtered were [16.882, 16.584] and [12.563, 12.014], respectively. The [train, validation] SSIM for pix2pix-raw and pix2pix-filtered were [0.617, 0.559] and [0.505, 0.448], respectively. A generator loss function value that is lower than a discriminator loss function value indicates that the generator can trick the discriminator into classifying its images as real more often than not. This occurred for both models, and suggested that each can generate realistic images. This was further confirmed by the SSIM and PSNR values. The higher SSIM and PSNR value of pix2pix-raw images, as compared to pix2pix-filtered images, suggested that its generated images were more similar to true retinal fundus images; this was likely due to the presence of choroidal blood vessels (**Figure 3**).



**Figure 4: Examples of real and generated retinal fundus images.** From left to right: a real retinal fundus image, a generated retinal fundus image from pix2pix-raw that uses raw vessel maps to create images with choroidal blood vessels patterns, and a generated retinal fundus image from pix2pix-filtered that uses filtered vessel maps to generate images without choroidal blood vessel patterns.

In general, experts were able to discern between real and generated images (Accuracy: 92.2%, **Table 1**). Images without choroidal blood vessel patterns (pix2pix-filtered) were identified as generated in 100% of cases. Some (16.7%) generated images that contained choroidal blood vessel patterns (pix2pix-raw) were classified as real images. This corroborates the difference in test set PSNR values between the two models. Nearly all (93.3%) of real images were identified as such. The  $\chi^2$  test statistically confirmed that, overall, experts could identify real versus generated images ( $\chi^2 \approx 64.019$ ;  $p \approx 1.254 \times 10^{-14}$ ). The Fisher's Exact Test confirmed this finding ( $p < 2.2 \times 10^{-16}$ ). Generated images with choroidal blood vessel patterns were, statistically, not more likely to be identified as real than those without ( $\chi^2$ :  $p \approx 0.062$ ; Fisher:  $p \approx 0.052$ ).

**Table 1: Expert majority determination of image type**

		Expert Majority Determination	
		Real	Generated
Image Type	Real Images	28	2
	pix2pix-raw	5	25
	pix2pix-filtered	0	30

Expert majority diagnoses for each image set are presented in **Table 2**. The majority diagnoses for real images were used as the ground truth. For normal images, experts diagnosed with accuracies of 92.3% and 100% on images generated by pix2pix-raw and pix2pix-filtered, respectively. For pre-plus images, experts had 91.7% accuracy on both pix2pix-raw and pix2pix-filtered images. Plus disease was diagnosed with 80% accuracy on both pix2pix-raw and pix2pix-filtered images. A Cochran-Mantel-Haenszel test confirmed that real, pix2pix-raw, and pix2pix-filtered images were not graded dissimilarly ( $p \approx 0.501$ ). This suggests that generated images, even those without choroidal blood vessel patterns, have the same diagnostic power as real images.

**Table 2: Expert majority diagnoses of real images versus expert majority diagnoses of generated images**

		Real Images			pix2pix-raw			pix2pix-filtered		
		Normal	Pre-Plus	Plus	Normal	Pre-Plus	Plus	Normal	Pre-Plus	Plus
Real Image Majority Diagnosis	Normal	13	0	0	12	1	0	13	0	0
	Pre-Plus	0	12	0	0	11	1	1	11	0
	Plus	0	0	5	0	1	4	0	1	4

To further investigate and confirm this finding, intergrader agreement of diagnoses and overall agreement of diagnoses between image sets were determined using weighted  $\kappa$  statistics for chance-adjusted agreement in ordinal diagnosis, using a well-known scale:  $[0, 0.20]$  = slight agreement,  $[0.21, 0.40]$  = fair agreement,  $[0.41, 0.60]$  = moderate agreement,  $[0.61, 0.80]$  = substantial agreement, and  $[0.81, 1.00]$  = near-perfect agreement (**Table 3**). For images generated by pix2pix-raw, individual expert diagnoses had substantial to near-perfect agreement ( $\kappa = [0.680, 0.880]$ ) to real images, and the majority diagnoses had near-perfect agreement ( $\kappa = 0.880$ ). For images generated by pix2pix-filtered, individual expert diagnoses had moderate to near-perfect agreement ( $\kappa = [0.498, 0.980]$ ) to real images, and the expert majority diagnoses had near-perfect agreement ( $\kappa = 0.902$ ).

**Table 3: Expert agreement of diagnoses between generated images and real images**

	Majority Diagnosis	Expert 1	Expert 2	Expert 3
pix2pix-raw	0.880	0.743	0.680	0.880
pix2pix-filtered	0.902	0.663	0.498	0.980

## Discussion

This study aimed to generate and evaluate synthetic retinal fundus images, for the diagnosis of plus disease in retinopathy of prematurity, by segmenting the vasculature of real retinal fundus images into grayscale vessel maps using a U-Net, and generating realistic color retinal fundus images from said vessel maps using pix2pix. There are two key findings: (1) images generated by pix2pix, regardless of whether choroidal blood vessel patterns were present, were easily identified by experts as generated, and (2) generated images retained information relevant for the detection of plus disease in retinopathy of prematurity. The Chi-squared test suggested that images were not realistic enough to deceive experts into believing they were real images ( $p \approx 1.254 \times 10^{-14}$ , **Table 1**). This was confirmed using a Fisher's Exact test ( $p < 2.2 \times 10^{-16}$ ). However, a Cochran-Mantel-Haenszel test showed that images were realistic enough to be diagnosed similarly to real images ( $p \approx 0.501$ , **Table 2**). This was further confirmed by measuring the agreement of individual and majority expert diagnoses across image sets using Cohen's kappa (**Table 3**). For real images, experts had near-perfect agreement with the diagnoses of the same images reconstructed by pix2pix-raw and pix2pix-filtered ( $\kappa = 0.880$ ,  $\kappa = 0.902$ ). These results suggest that, although images may be recognized as generated, these models retain relevant information that is required to reconstruct retinal fundus images for the diagnosis of plus disease in retinopathy of prematurity.

This work has many important implications. First, it confirms that, at least for plus disease diagnosis in ROP, pix2pix-generated images are of high enough quality and fidelity for expert physicians to form accurate diagnoses. This is important, as numeric metrics and subjective layperson evaluations are often used to evaluate the quality and realism of generated images.<sup>5,7,9,10</sup> However, because the goals of such systems are often to generate images for synthetic datasets, training physicians, or diagnosis from image reconstructions, the ability of physicians to form diagnoses from generated images should be evaluated prior to implementation. It should also be noted that clinical

findings unrelated to plus disease diagnosis that were present in original retinal fundus images, such as hemorrhages, were not present in reconstructions. This is likely because the U-Net was not trained to segment retinal hemorrhages, so pix2pix was unaware of their existence. While not detrimental for the diagnosis of plus disease, it serves as a warning to those looking to generate images of highly-complex diseases where rare clinical findings may be highly-relevant for a given diagnosis. In essence, although an image may appear real or diagnosable, it may be lacking pertinent information that was present in the original image.

Second, retinal scans are listed as protected health information (PHI) according to the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule; however, they are de-identified via the Expert Determination method (§ 164.514(b)(2)).<sup>20,21</sup> This method essentially states that the risk of re-identification using retinal fundus images is negligible; however, the European Union's General Data Protection Regulation (GDPR) currently informs that this method is not sufficient for de-identification. It is conceivable that the pix2pix methods we have trained could be used to further de-identify retinal fundus images. From a purely observational standpoint, it was found that generated images were often pigmented differently than original images (**Figure 3** and **Figure 4**). Second, generated images without choroidal blood vessel patterns were just as diagnosable as real images. Finally, other clinical findings present in retinal fundus images, such as hemorrhages, may be specific to unique diagnoses and increase the identifiability of images. As mentioned, in the i-ROP dataset, there were a limited number of images with these features; therefore, the few retinal fundus images that did have these highly-identifiable features were reconstructed without them. Although this method would not fully de-identify images, it could further reduce the risk of re-identification by removing highly-identifiable features while still providing physicians with the information required to form accurate diagnoses. One could argue that the retinal vessel maps used to generate retinal fundus images are computationally easier to generate and modify, and might be considered further de-identified. However, physicians are not formally trained to form diagnoses from these types of images, and the accuracy and reliability of said diagnoses could suffer. Nonetheless, it is one of a few interesting future directions for this work.

In addition, the task of increasing the fidelity of generated retinal fundus images from pix2pix-raw will be pursued, using methods such as pix2pixHD, with the goal to trick experts into always classifying generated images as real.<sup>22</sup> This may have applications for scenarios where one wishes to modify the retinal vasculature and generate a highly-realistic retinal fundus image. Second, the level of U-Net vessel map filtering for pix2pix-filtered will be increased to evaluate what threshold is required for experts to adequately diagnose ROP images. This will inform at what point images begin to resemble retinal vessel maps too closely, and diagnostic accuracy begins to drop. Finally, this evaluation method will be applied to other ophthalmic diseases, such as diabetic retinopathy.

## **Conclusion**

We have implemented the pix2pix generative adversarial network for the generation of retinal fundus images in retinopathy of prematurity. We have trained models that can generate somewhat realistic retinal fundus images from retinal vessel maps. While these generated images can be identified as such, they still have the same diagnostic power as real images; both sets of images were easily diagnosed for the presence of plus disease by retinopathy of prematurity experts. This is important, as previous studies have not formally evaluated the ability of physicians to form diagnoses from generated medical images.

## **Acknowledgements**

This study was supported by grants T15LM007088, R01EY19474, K12EY027720, and P30EY10572 from the National Institutes of Health (Bethesda, MD), the Research to Prevent Blindness Medical Student Fellowship (New York, NY), and by unrestricted departmental funding from Research to Prevent Blindness (New York, NY).

## **Conflicts of Interest**

No financial conflicts of interest. Michael F. Chiang is an unpaid member of the Scientific Advisory Board for Clarity Medical Systems (Pleasanton, CA), and a Consultant for Novartis (Basel, Switzerland).

## References

1. Wang H, Peng H, Chang Y, Liang D. A survey of GPU-based acceleration techniques in MRI reconstructions. *Quant Imaging Med Surg*. 2018;8(2):196-208.
2. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.
3. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks. June 2014. <http://arxiv.org/abs/1406.2661>. Accessed March 4, 2020.
4. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-Image Translation with Conditional Adversarial Networks. November 2016. <http://arxiv.org/abs/1611.07004>. Accessed March 3, 2020.
5. Andreini P, Bonechi S, Bianchini M, Mecocci A, Scarselli F, Sodi A. A Two Stage GAN for High Resolution Retinal Image Generation and Segmentation. July 2019. <http://arxiv.org/abs/1907.12296>. Accessed March 23, 2020.
6. Beers A, Brown J, Chang K, et al. High-resolution medical image synthesis using progressively grown generative adversarial networks. May 2018. <http://arxiv.org/abs/1805.03144>. Accessed March 23, 2020.
7. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. *Med Image Anal*. 2019;58:101552.
8. Costa P, Galdran A, Meyer MI, et al. Towards Adversarial Retinal Image Synthesis. January 2017. <http://arxiv.org/abs/1701.08974>. Accessed March 23, 2020.
9. Yu Z, Xiang Q, Meng J, Kou C, Ren Q, Lu Y. Retinal image synthesis from multiple-landmarks input with generative adversarial networks. *Biomed Eng Online*. 2019;18. doi:10.1186/s12938-019-0682-x
10. End-to-End Adversarial Retinal Image Synthesis - IEEE Journals & Magazine. <https://ieeexplore.ieee.org/document/8055572>. Accessed March 23, 2020.
11. Retinopathy of Prematurity | National Eye Institute. <https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/retinopathy-prematurity>. Accessed March 3, 2020.
12. The International Classification of Retinopathy of Prematurity Revisited. *Archives of Ophthalmology*. 2005;123(7):991. doi:10.1001/archophth.123.7.991
13. Brown JM, Campbell JP, Beers A, et al. Automated Diagnosis of Plus Disease in Retinopathy of Prematurity Using Deep Convolutional Neural Networks. *JAMA Ophthalmol*. 2018;136(7):803-810.
14. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. May 2015. <http://arxiv.org/abs/1505.04597>. Accessed March 4, 2020.
15. Paszke A, Gross S, Massa F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. December 2019. <http://arxiv.org/abs/1912.01703>. Accessed March 4, 2020.
16. Image-to-Image Translation with Conditional Adversarial Networks. <https://phillipi.github.io/pix2pix/>. Accessed March 4, 2020.
17. Welstead S. Fractal and Wavelet Image Compression Techniques. 1999. doi:10.1117/3.353798
18. Image quality assessment: from error visibility to structural similarity - IEEE Journals & Magazine. <https://ieeexplore.ieee.org/document/1284395>. Accessed March 17, 2020.
19. Kalpathy-Cramer J, Campbell JP, Erdogmus D, et al. Plus Disease in Retinopathy of Prematurity: Improving

Diagnosis by Ranking Disease Severity and Using Quantitative Image Analysis. *Ophthalmology*. 2016;123(11):2345-2351.

20. Office for Civil Rights (OCR). Summary of the HIPAA Privacy Rule. HHS.gov. <https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html>. Published July 26, 2013. Accessed March 3, 2020.
21. Office for Civil Rights (OCR). Methods for De-identification of PHI. HHS.gov. <https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html>. Published November 6, 2015. Accessed March 3, 2020.
22. Wang T-C, Liu M-Y, Zhu J-Y, Tao A, Kautz J, Catanzaro B. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. November 2017. <http://arxiv.org/abs/1711.11585>. Accessed March 23, 2020.