tvst

Article

Assessment of Generative Adversarial Networks for Synthetic Anterior Segment Optical Coherence Tomography Images in Closed-Angle Detection

Ce Zheng[1], Fang Bian[2,3], Luo Li[2], Xiaolin Xie[2], Hui Liu[4], Jianheng Liang[4], Xu Chen[4,5], Zilei Wang[1], Tong Qiao[1], Jianlong Yang[6], and Mingzhi Zhang[2]

[1] Department of Ophthalmology, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai, China
[2] Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong, Shantou University Medical College, Shantou, Guangdong, China
[3] Department of Ophthalmology, Deyang People's Hospital, Sichuan, China
[4] Aier School of Ophthalmology, Central South University, Changsha, Hunan, China
[5] Department of Ophthalmology, Shanghai Aier Eye Hospital, Shanghai, China
[6] Ningbo Institute of Industrial Technology, Chinese Academy of Sciences, Ningbo, Zhejiang, China

Correspondence: Mingzhi Zhang, Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong, Shantou University Medical College, Shantou 515000, China. e-mail: zmz0754@126.com

Received: June 20, 2020
Accepted: March 8, 2021
Published: April 30, 2021

Keywords: anterior segment optical coherence tomography; closed-angle; deep learning; generative adversarial networks

Citation: Zheng C, Bian F, Li L, Xie X, Liu H, Liang J, Chen X, Wang Z, Qiao T, Yang J, Zhang M. Assessment of generative adversarial networks for synthetic anterior segment optical coherence tomography images in closed-angle detection. Transl Vis Sci Technol. 2021;10(4):34, https://doi.org/10.1167/tvst.10.4.34

Purpose: To develop generative adversarial networks (GANs) that synthesize realistic anterior segment optical coherence tomography (AS-OCT) images and evaluate deep learning (DL) models that are trained on real and synthetic datasets for detecting angle closure.

Methods: The GAN architecture was adopted and trained on the dataset with AS-OCT images collected from the Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong, synthesizing open- and closed-angle AS-OCT images. A visual Turing test with two glaucoma specialists was performed to assess the image quality of real and synthetic images. DL models, trained on either real or synthetic datasets, were developed. Using the clinicians' grading of the AS-OCT images as the reference standard, we compared the diagnostic performance of open-angle vs. closed-angle detection of DL models and the AS-OCT parameter, defined as a trabecular-iris space area 750 µm anterior to the scleral spur (TISA750), in a small independent validation dataset.

Results: The GAN training included 28,643 AS-OCT anterior chamber angle (ACA) images. The real and synthetic datasets for DL model training have an equal distribution of open- and closed-angle images (all with 10,000 images each). The independent validation dataset included 238 open-angle and 243 closed-angle AS-OCT ACA images. The image quality of real versus synthetic AS-OCT images was similar, as assessed by the two glaucoma specialists, except for the scleral spur visibility. For the independent validation dataset, both DL models achieved higher areas under the curve compared with TISA750. Two DL models had areas under the curve of 0.97 (95% confidence interval, 0.96–0.99) and 0.94 (95% confidence interval, 0.92–0.96).

Conclusions: The GAN synthetic AS-OCT images appeared to be of good quality, according to the glaucoma specialists. The DL models, trained on all-synthetic AS-OCT images, can achieve high diagnostic performance.

Translational Relevance: The GANs can generate realistic AS-OCT images, which can also be used to train DL models.

Introduction

Primary angle-closure glaucoma (PACG) is a major cause of visual impairment in Asia, with a prevalence of between 1.0% and 1.4% in Asians 40 years and older.[1] In China, PACG may blind ten times more people than primary open-angle glaucoma.[2] Moreover, it is estimated that the prevalence of PACG will reach 32.04 million in 2040.[3] Anterior segment optical coherence tomography (AS-OCT), which uses low-coherence interferometry laser to obtain a cross-sectional image and provide objective assessment,[4] is one of the most common imaging modalities for anterior chamber angle (ACA) assessment in current clinical practice. With advances in deep learning (DL), it is now possible to automatically detect angle closure directly from AS-OCT images. Recently, Xu et al.[5] demonstrated highly accurate DL classifiers for detecting gonioscopic angle closure with an area under the receiver operating characteristic curve (AUC) of 0.933 from a population-based study. However, generalizable DL models require large amounts of diverse and well-labeled data generated by experts, which is time consuming and suffers from inherent inter-rater variability.[6] Training on 8270 ACA images using the Visante AS-OCT (Carl Zeiss Meditec, Jena, Germany), Fu et al.[7] reported a DL system that achieved an AUC of 0.96 using iris-trabecular contact (ITC) as the reference standard. In another study, the same author also developed DL algorithms to detect ITC using two different AS-OCT modalities (Visante and Cirrus; Carl Zeiss Meditec) with nearly 10,000 ACA images.

Generative adversarial networks (GANs),[8] which are inspired by game theory for training a model in an adversarial process, offer a novel method to generate new medical images. Such an approach has several advantages. First, it helps overcome the privacy issues related to biomedical image data, which is a kind of personally identifiable information as defined by the National Institute of Standard and Technology.[9] Second, several studies have already demonstrated that generated medical images can improve the performance of DL algorithms for medical image classification.[10–12]

In a previous study, we proposed a GAN approach to generate realistic OCT images to serve as training datasets for DL algorithms and education images for retinal specialists.[13] We showed that DL algorithms trained with only generated images achieved performance nearly comparable to the results obtained from training on real images. In the present work, we first demonstrate that a similar technique can also generate realistic AS-OCT images. We further assess whether GAN-generated AS-OCT images can satisfactorily serve as training datasets to develop DL algorithms for the detection of angle closure.

Methods

This study was conducted according to the tenets of the Declaration of Helsinki and was approved by the institutional review board of the Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong (JSIEC) (identifier, 2018RY029-E01). Written informed consent was obtained from all the subjects.

Datasets

Chinese subjects, including patients with angle closure and normal subjects with open angle, were recruited from JSIEC between September 2014 and May 2016. All subjects were asked about their medical and ophthalmic history and underwent a standardized ophthalmic examination including best-corrected visual acuity, refraction, slit-lamp biomicroscopy, intraocular pressure (IOP) measurement by Goldmann applanation tonometry, and fundus examination, as well as biometry (Zeiss IOLMaster; Carl Zeiss Meditec). Gonioscopy was performed in a dark room by a glaucoma specialist (CZ) using a Goldmann two-mirror lens (Haag-Streit AG, Bern, Switzerland) at 16× magnification. The modified Shaffer grading system (grade 0 = no structures visible; grade 1 = Schwalbe's line visible; grade 2 = anterior trabecular meshwork visible; grade 3 = posterior trabecular meshwork or scleral spur visible; grade 4 = ciliary body visible) was used to evaluate the angle on gonioscopy.[14] To establish the presence and degree of peripheral anterior synechiae, indentation gonioscopy was carried out using a Sussman four-mirror lens (Ocular Instruments, Inc., Bellevue, WA) under dark conditions. During gonioscopy, the slit light was reduced to 1 mm to avoid falling on the pupil. The gonioscopy lens could be tilted to exam the angle over the convexity of the iris. Based on the definition from the American Academy of Ophthalmology Preferred Practice Patterns,[15] angle closure was classified into one of the following subtypes: (1) Primary angle-closure suspect is defined in eyes that have two quadrants or more of ITC without elevated IOP or optic nerve damage. (2) Primary angle closure is defined in eyes that meet primary angle-closure suspect criteria and also have the presence of peripheral anterior synechiae and/or elevated IOP (>21 mm Hg); primary angle-closure glaucoma (PACG) is defined as primary angle

closure with glaucomatous optic neuropathy. (3) Acute primary-angle closure (APAC) is defined as presenting two or more symptoms (ocular or periocular pain; nausea or vomiting, or both; and an antecedent history of intermittent blurring of vision), IOP > 28 mm Hg, and closed angles in four quadrants found on gonioscopic examination. Inclusion criteria for the normal open-angle group were (1) age > 20 years old; (2) best-corrected visual acuity of 0.3 logMAR or better with spherical refraction between –6.0 and 6.0 diopters; (3) IOP ≤ 21 mm Hg; (4) no history of glaucoma, retinal disease, diabetic mellitus, or significant senile cataract that could affect the results of AS-OCT examinations; or (5) normal visual field, which was defined as a mean deviation and pattern standard deviation within 95% confidence intervals (CIs) and a glaucoma hemifield test result within normal limits. Exclusion criteria included the following: (1) eyes with corneal abnormalities, such as edema, pterygium, and other degenerative changes; (2) eyes with previous laser peripheral iridotomy or iridoplasty; or (3) eyes with previous intraocular surgery, such as cataract surgery, retinal surgery, or trabeculectomy.

ACA images were obtained with Casia SS-1000 AS-OCT (Tomey Corporation, Nagoya, Japan) and collected for GAN training. The details of the Casia AS-OCT modality have been described previously. In brief, Casia AS-OCT has a swept laser source and a scan speed of 30,000 A-scans per second. For image acquisition, a seated subject was directed to an internal fixation under dark conditions, and upper and lower eyelids were gently retracted for imaging of superior and inferior quadrants. A total of 128 two-dimensional cross-sectional AS-OCT images were acquired per eye using the Casia SS-1000 AS-OCT. All AS-OCT images were evaluated both qualitatively and quantitatively. For the clinicians' qualitative evaluation, a closed angle in AS-OCT images was defined as at least substantial ITC between the peripheral iris and at least one-third of the trabecular meshwork anterior to the scleral spur.[7] Interobserver variability among graders was good, with an unweighted k value of more than 0.75 (50 AS-OCT images with open angle and 50 AS-OCT images with closed angle randomly selected from the training dataset). Customized software (Anterior Segment Analysis Program) was used to measure TISA750, which is an AS-OCT biometric parameter[16] and defined as a trabecular-iris space area at 750 um anterior to the scleral spur.

We also collected an independent validation dataset, which included 238 open-angle and 243 angle-closure AS-OCT ACA images from the same center between March 2019 and January 2020 to evaluate the discriminatory ability of the DL algorithm and AS-OCT parameters. The inclusion and exclusion criteria were identical for both the training and external validation datasets, except that only one horizontal AS-OCT image of each eligible subject was considered for the study in the testing dataset.

Development of GAN Synthetic AS-OCT Images

We have reported the details regarding GAN synthetic OCT imaging in our previous study. Briefly, a progressively grown generative adversarial network (PGGAN) is implemented by incrementally increasing the size of the synthetic images (e.g., 8 × 8, 16 × 16) until the desired output size (in this study, 256 × 256 pixels) is met (see Fig. 2). A PGGAN is an extension to a GAN wherein a generator network (G) is trained to initially synthesize low-resolution synthetic images (4 × 4 pixels), which are then fed to a discriminator network (D) that distinguishes these synthetic images from real images. With adversarial training between the G and the D, a GAN can generate realistic synthetic images at a high resolution. Two separate PGGANs (open-angle GAN and closed-angle GAN) were developed to generate two separate synthetic AS-OCT datasets: one for open-angle AS-OCT ACA images and one for closed-angle AS-OCT ACA images. For implementation of the PGGAN architectures, we followed the scheme suggested by Karras et al.[17] and added an upsampling layer and a pair of convolutional layers to both the D and G networks with each phase. An upsampling layer involves two convolutional layers with 3 × 3 and 4 × 4 filters and a Leaky ReLU activation function with a slope of 0.2. Average pooling is used for downsampling. The PGGAN model was implemented with the TensorFlow 2.1.0 framework and Keras 2.2.4 API. All training processes were performed on an NVIDIA (Santa Clara, CA) GTX 1080 Ti 12-GB GPU with an Intel (Santa Clara, CA) Core i7-2700K processor, 4.6-GHz central processing unit, and 128 GB RAM.

Evaluation of GAN Synthetic AS-OCT Images

We performed a visual Turing test to qualitatively evaluate the image quality of synthetic AS-OCT ACA images.[18,19] Two glaucoma specialists manually assessed the image quality of a subset of 100 AS-OCT ACA images with 256 × 256 square pixels, comprised of 50 real images and 50 synthetic images. Neither glaucoma specialist was told that the subset images were composed of a mixture of real and synthetic images. Image quality was evaluated as follows: (1)

good visibility of the scleral spurs (defined as the point where there was a change in curvature of the inner surface of the angle wall),[20] (2) the presence of continuity in the anterior segment structures, and (3) the absence of motion artifacts.[21] During the Turing test, two glaucoma specialists (both had more than 10 years of clinical experience) were blinded to image information (real or synthetic image) and independently graded the image quality. To visualize the high-dimensional properties between synthetic and real AS-OCT ACA images, we qualitatively compared them with images generated by the PGGAN and the real images by using *t*-distributed stochastic neighbor embedding (t-SNE) plots.[22] t-SNE plots take high-dimensional datasets and reduce them to low-dimensional graphs that retain most of the original information.

We also assessed whether synthetic OCT images can be used as a training dataset for DL models. A similar method had been reported by us previously.[13] Briefly, we modified a Google (Menlo Park, CA) Inception V3 architecture with weights pretrained on ImageNet using the transfer learning technique.[23,24] DL_Model_R trained exclusively with real images (n = 20,000, randomly selected from the GAN training dataset with all 28,643 AS-OCT ACA images), and DL_Model_S trained only with synthetic images (n = 20,000). The classification performance was then compared among DL_Model_R, DL_Model_S, and AS-OCT biometric parameters (TISA750) in the independent validation dataset.

Statistics

The statistical analysis was done in terms of the receiver operating characteristic curve, accuracy, sensitivity, and specificity with two-sided 95% CIs. The AUC was used to assess the performance of the DL models and TISA750 with regard to detecting angle closure in AS-OCT images. Sensitivity and specificity were calculated for the optimum cutoff values obtained from receiver operating characteristic curves. We calculated accuracy, sensitivity, and specificity using the following formulas:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \quad (1)$$

$$\text{Specifficity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

TP, TN, FP, and FN are true positive, true negative, false positive, and false negative, respectively. The χ^2 test was used to compare categorical data.

Results

We evaluated 30,782 AS-OCT ACA images of 1439 eyes from 894 subjects who were previously enrolled for the AS-OCT imaging at JSIEC. We excluded 2139 ACA images (92 eyes from 57 subjects) due to poor-quality images (eyelid artifact, corrupt images, or motion artifact) (Fig. 1). To eliminate potential sources of bias, we further retrained the training dataset with equal distributions of open- and closed-angle images. The final AS-OCT training dataset for developing the PGGAN model consisted of 20,000 AS-OCT ACA images (1347 eyes from 837 subjects with 335 closed-angle eyes and 503 open-angle eyes with gonioscopy diagnosed) (Table 1). The subjects' mean age was 63.01 ± 8.46 years, and 367 subjects (43.8%) were male. The average of TISA750 was significantly smaller in the closed-angle group than in the open-angle group in all datasets (all with $P < .001$).

Figure 2 illustrates the training progress of the schematic model of a PGGAN. The PGGAN starts at low resolution (4 × 4-pixel images), doubling to 8 × 8, 16 × 16, and so on until the desired output resolution is reached. A t-SNE visualization of the low-dimensional features associated with real and synthetic AS-OCT ACA images is shown in Figure 3. From t-SNE analysis, it is possible to see that the low-dimensional features of the synthetic images generated by the GAN architecture are different from the real images. The results of real and synthetic AS-OCT ACA image quality being graded by the two glaucoma specialists are shown in Table 2. Both real and synthetic AS-OCT ACA images have the approximately same quality based on the input of the two glaucoma specialists, except for scleral spur visibility. For both glaucoma specialists, the frequency of detectable scleral spurs in real AS-OCT images was higher than that in the synthetic AS-OCT images (all with $P < 0.01$).

Figure 4 shows the AUCs of two DL models and TISA750 testing in the validation dataset using the clinicians' qualitative evaluation as reference. Both DL models achieved higher AUCs compared to TISA750. TISA750 had an AUC of 0.88 (95% CI, 0.81–0.91) with a cut-off value of 0.019 mm^2. DL_Model_R (trained on real AS-OCT ACA images) obtained the best performance, with an AUC of 0.97 (95% CI, 0.96–0.99). Although the AUC of DL_Model_S (trained on synthetic AS-OCT ACA images) indicated

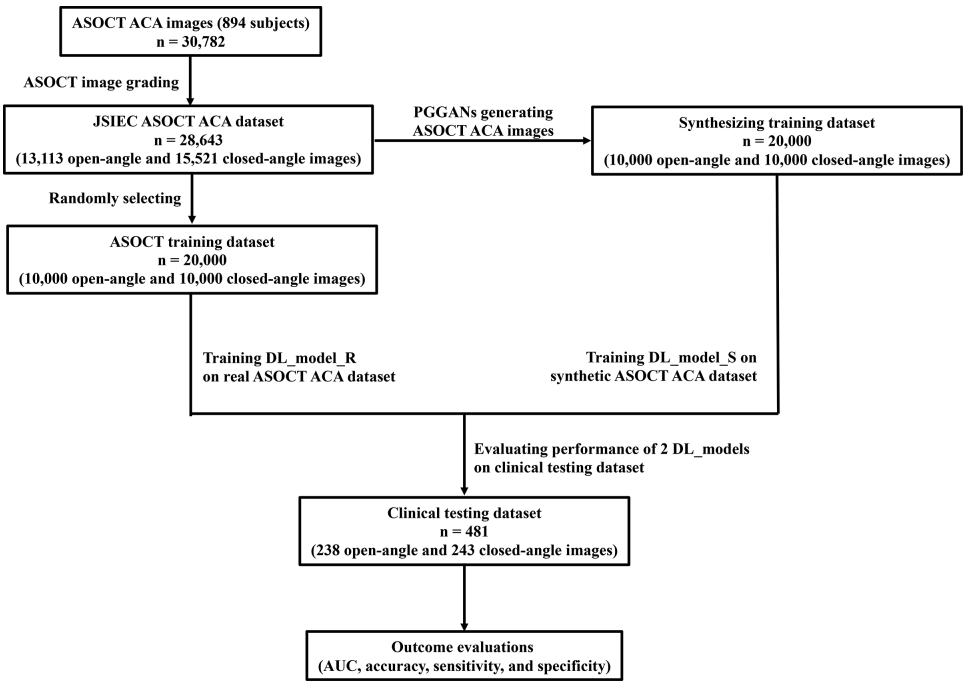

Figure 1. Schematic of AS-OCT anterior chamber angle image grading and using synthetic AS-OCT images in closed-angle classification.

Table 1. Baseline Clinical and AS-OCT Biometric Characteristics of the JSIEC Development and Local Clinical Validation Datasets

Characteristics	Development Dataset	Local Clinical Validation Dataset
Participants, *n*	837	481
ACA images (open- vs. closed-angle), *n*	13,000 vs. 13,000	238 vs. 243
Age (y), mean ± SD	63.01 ± 8.46	65.71 ± 9.03
Male, %	43.80	36.38
Right eye, %	48.50	54.83

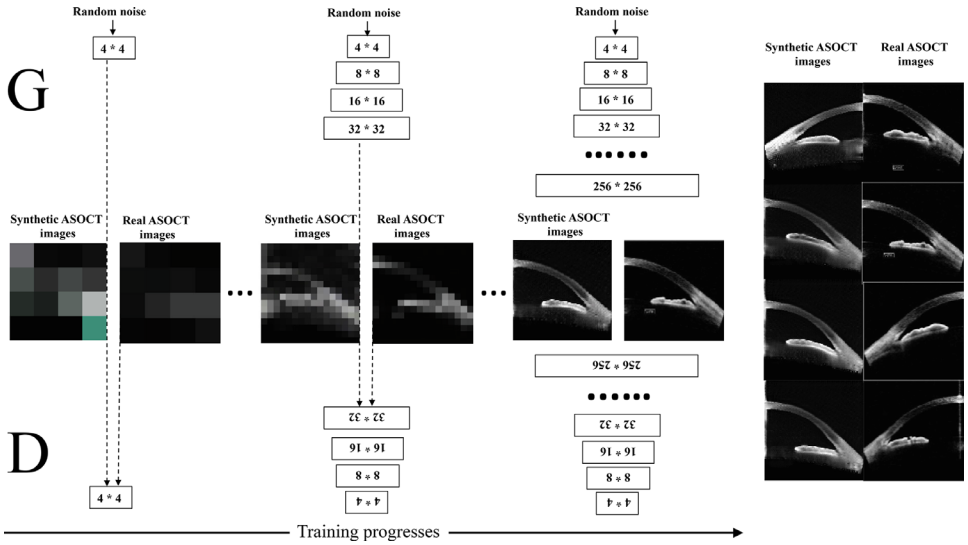

Figure 2. Schematic for generating synthetic AS-OCT images by PGGANs. The G and D networks are simultaneously trained. Using random noise as input, the G networks create synthetic AS-OCT ACA images. The real images are then fed to the D networks, which output a probability of being real or fake. The PGGAN architecture trains a single GAN in a stepwise fashion from 4 to 256 square pixels, respectively.

Table 2. AS-OCT ACT Image Quality (Synthetic vs. Real) As Graded by Two Glaucoma Specialists

	Synthetic, *n* (%)	Real, *n* (%)	*P*
Glaucoma specialist 1			
Visibility of scleral spurs	32 (64)	45 (90)	0.005
Continuity in anterior segment structures	48 (96)	47 (94)	0.646
Absence of motion artifacts	46 (92)	49 (98)	0.169
Glaucoma specialist 2			
Visibility of scleral spurs	35 (70)	47 (94)	0.002
Continuity in anterior segment structures	45 (90)	49 (98)	0.092
Absence of motion artifacts	50 (100)	49 (98)	0.315

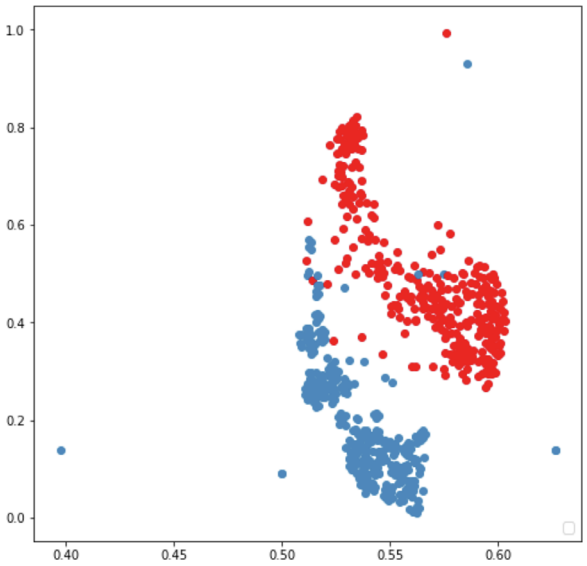

Figure 3. A t-SNE visualization of the features associated with real and synthetic AS-OCT ACA images. *Red* and *blue dots* indicate synthetic and real features, respectively.

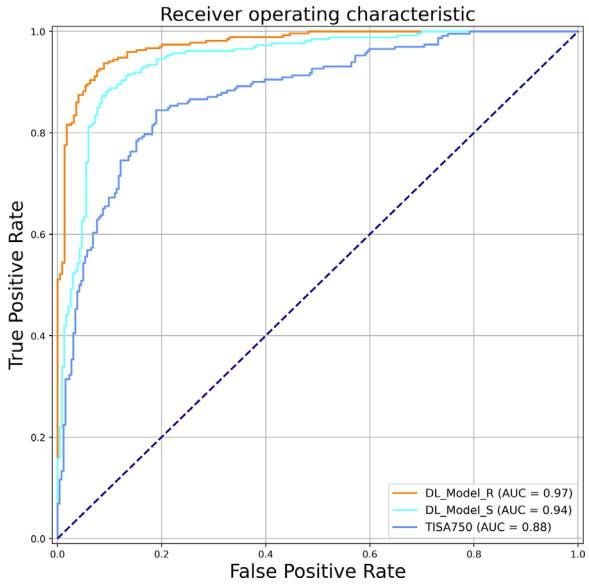

Figure 4. The average AUCs of two DL models and AS-OCT parameter (TISA750) testing in the independent validation dataset. DL_Model_R is a deep learning model trained on real AS-OCT ACA images, and DL_Model_S is a deep learning model trained on synthetic AS-OCT ACA images.

moderately lower performance compared with that of DL_Model_R, the performance was still considered to be good, with an AUC of 0.94 (95% CI, 0.92–0.96).

Discussion

In the current study, we have introduced a PGGAN approach for generating high-resolution AS-OCT ACA images; we evaluated image quality graded by glaucoma specialists and trained DL models to detect angle closure using synthetic AS-OCT ACA images. Our results demonstrated that, although glaucoma specialists can identify synthetic AS-OCT ACA images due to lower visibility of the scleral spurs and anterior segment structures, excellent performance can be achieved to detect angle closure in AS-OCT ACA images by DL models when we used synthetic images as data augmentation techniques.

Deep learning, such as deep neural networks, has gained rapid success in medical imaging.[25–27] GANs are a type of deep neural network aimed at generating new images.[28] The application of GANs includes data augmentation, image reconstruction, image segmentation, and image transformation between different modalities.[29–31] Using PGGANs, Beers et al.[32] demonstrated that this approach could produce realistic medical images in various domains, including fundus photographs with retinopathy of prematurity. Several other studies have also showed that it is promising to generate synthetic medical images for data augmentation to improve classification performance with limited data[33–35] in different medical imaging domains, such as neuronal imaging, stain-free cancer cell imaging, or CT imaging of liver lesions. In our previous study, we reported that GANs were able to

synthesize realistic high-resolution OCT images and achieve a high AUC of 0.98 for screening urgent referable retinal diseases, such as choroidal neovascularization or diabetic macular edema.[13] In the current study, our DL model that trained on synthetic OCT images achieved a similar AUC of 0.94, which was comparable with that of DL models trained in all-real OCT images, such as those reported by Xu et al.[5] (4036 AS-OCT images) and Fu et al.[7] (8270 AS-OCT ACA images). As different reference standards were used, our study cannot be directly compared to those of Xu et al. and Fu et al. It should be noted that a small validation dataset makes it challenging to interpret a small difference in model performance, and the distribution of open- and closed-angle images was not the same in the two training datasets, which could also introduce training bias. But, our results show the potential application of this technique to image augmentation for DL model training.

How to evaluate GAN synthetic images is still controversial. Unlike other DL models trained with a loss function until convergence, both the G and D models are trained together to maintain an equilibrium in GANs. As such, there is no objective way to assess the relative or absolute quality of the model from loss alone.[36] Schlegl et al.[19] proposed a visual Turing test to quantify the quality of the GANs generated retinal images by two clinical retina experts. Based on a GAN model similar to that described in the current study, Burlina et al.[18] suggested that GAN-synthesized macular images appeared to be realistic to retinal specialists. In the present study, we further evaluated synthetic image quality with stricter criteria proposed by Lee et al.[21] Our results demonstrate that synthetic AS-OCT ACA image quality is comparable to the real image except for lower visibility of the scleral spurs. Scleral spur location represents a critical anatomical landmark for imaging the anterior chamber angle, as it is a landmark for quantitative measurements of the ACA. Using the Visante AS-OCT, Sakata et al.[20] reported that scleral spur location could be determined in 72% of the AS-OCT images. Our results are encouraging in that GANs still generated high-quality AS-OCT images, although the frequency of detectable scleral spurs in real AS-OCT images is higher than that in synthetic AS-OCT images (90% and 94% vs. 64% and 70%, respectively; all $P < 0.01$). There are two possible explanations for the lower visibility of sclera spurs in synthetic images. First, synthesized AS-OCT ACA images have a lower resolution than Casia AS-OCT images. Second, a GAN itself can produce artifacts that eliminate details present in the original image.[37] Another possible concern is generating images too similar to the original training images, or even just copying and pasting all training images. We used t-SNE, a nonlinear technique for embedding high-dimensional data in a low-dimensional space suited for visualization to address this issue. Our results suggest that GAN synthetic images are still distinct from original training images.

In the current study, we focused on assessing GANs dedicated to generating AS-OCT images with open or closed angles rather than evaluating the diagnostic performance of hand-crafted features (HCFs) such as TISA750. Using a custom program and HCFs, Narayanaswamy et al.[38] reported that the AUCs for TISA750 in the nasal and temporal quadrants were 0.87 and 0.88, respectively. Such methods are time consuming, and the results may be compromised due to interobserver variability. Our results indicate that the DL models significantly outperformed HCFs for the detection of angle closure from AS-OCT images. However, the black-box behavior of DL models hinders their application in clinical settings.[39] Most clinicians still prefer HCFs, as they can reflect expert knowledge. Further study, such as determining HCFs from GANs synthetic images and comparing the predicted features to the real features, would help clinicians understand DL model decisions.

Our study has some limitations. First, we synthesized only AS-OCT ACA images with 256 × 256 square pixels, similar to that of Visante AS-OCT images (300 × 300 square pixels) but lower than that of Casia AS-OCT images. It is possible for PGGANs to generate higher resolutions (e.g., 1024 × 1024 or above). Future work will involve evaluations at higher resolutions using similar PGGAN models to evaluate different angle-closure mechanisms. Second, although we assessed the generalizability of the DL model by testing it within independent external datasets collected from the same center, several studies have reported different angle closure mechanisms in diverse populations.[40] The application of the PGGAN technique requires further validation in multicenter and multiethnic trials. Third, only two glaucoma specialists were enrolled to perform the visual Turing testing in the current study. Further study is warranted to investigate the interrater reliability of multiple human graders. Finally, all AS-OCT images were manually graded, which may lead to interobserver variability.

In summary, we have suggested a GAN approach for obtaining good-quality synthetic AS-OCT images that address the problem of a small training dataset for detecting angle closure. Future work is needed to extend our work to other ophthalmological domains that can benefit from the synthesis of new images for improved DL model training and classification results.

Acknowledgments

Supported by the National Natural Science Foundation of China (81371010) and Clinical Research Funds of Shantou University Medical College (2014).

Disclosure: **C. Zheng,** None; **F. Bian,** None; **L. Li,** None; **X. Xie,** None; **H. Liu,** None; **J. Liang,** None; **X. Chen,** None; **Z. Wang,** None; **T. Qiao,** None; **J. Yang,** None; **M. Zhang,** None

References

1. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006;90(3):262–267.
2. Foster PJ, Johnson GJ. Glaucoma in China: how big is the problem? *Br J Ophthalmol*. 2001;85(11):1277–1282.
3. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng C-Y. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014;121(11):2081–2090.
4. Nolan WP, See JL, Chew PT, et al. Detection of primary angle closure using anterior segment optical coherence tomography in Asian eyes. *Ophthalmology*. 2007;114(1):33–39.
5. Xu BY, Chiang M, Chaudhary S, Kulkarni S, Pardeshi AA, Varma R. Deep learning classifiers for automated detection of gonioscopic angle closure based on anterior segment OCT images. *Am J Ophthalmol*. 2019;208:273–280.
6. Ting DSW, Liu Y, Burlina P, Xu X, Bressler NM, Wong TY. AI for medical imaging goes deep. *Nat Med*. 2018;24(5):539–540.
7. Fu HZ, Mani B, Xu YW, et al. A deep learning system for automated angle-closure detection in anterior segment optical coherence tomography images. *Am J Ophthalmol*. 2019;203:37–45.
8. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Adv Neural Inf Process Syst*. 2014;2:2672–2680.
9. McCallister E, Grance T, Scarfone KA. *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII): Recommendations of the National Institute of Standards and Technology*. Gaithersburg, MD: National Institute of Standards and Technology; 2011:800–122.
10. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*. 2018;321:321–331.
11. Sandfort V, Yan K, Pickhardt PJ, Summers RM. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Sci Rep*. 2019;9(1):16884.
12. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal*. 2019;58:101552.
13. Zheng C, Xie XL, Zhou K, et al. Assessment of generative adversarial networks model for synthetic optical coherence tomography images of retinal disorders. *Transl Vis Sci Technol*. 2020;9(2):29.
14. Foster PJ, Buhrmann R, Quigley HA, Johnson GJ. The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol*. 2002;86(2):238–242.
15. Prum BE, Herndon LW, Moroi SE, et al. Primary Angle Closure Preferred Practice Pattern guidelines. *Ophthalmology*. 2016;123(1):P1–P40.
16. Zheng C, Cheung CY, Aung T, et al. In vivo analysis of vectors involved in pupil constriction in Chinese subjects with angle closure. *Invest Ophthalmol Visual Sci*. 2012;53(11):6756–6762.
17. Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of GANs for improved quality, stability, and variation. Paper presented at the Sixth International Conference on Learning Representations (ICLR 2018), Vancouver, Canada, April 30–May 3, 2018.
18. Burlina PM, Joshi N, Pacheco KD, Liu TYA, Bressler NM. Assessment of deep generative models for high-resolution synthetic retinal image generation of age-related macular degeneration. *JAMA Ophthalmol*. 2019;137(3):258–264.
19. Schlegl T, Seeböck P, Waldstein SM, Langs G, Schmidt-Erfurth U. f-AnoGAN: Fast Unsupervised Anomaly Detection with Generative Adversarial Networks. *Med Image Anal*. 2019;54:30–44.
20. Sakata LM, Lavanya R, Friedman DS, et al. Assessment of the scleral spur in anterior segment optical coherence tomography images. *Arch Ophthalmol*. 2008;126(2):181–185.
21. Lee RY, Kasuga T, Cui QN, et al. Association between baseline iris thickness and prophylactic laser peripheral iridotomy outcomes in primary angle-closure suspects. *Ophthalmology*. 2014;121(6):1194–1202.
22. Laurens VDM, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res*. 2008;9(2605):2579–2605.

23. Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2016:27–30.
24. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Paper presented at the Third International Conference on Learning Representations (ICLR 2015), San Diego, CA, May 7–May 9, 2018.
25. James M, Brown J, Peter C, et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. *JAMA Ophthalmol*. 2018;136(7):803–810.
26. Burlina P, Pacheco KD, Joshi N, Freund DE, Bressler NM. Comparing humans and deep learning performance for grading AMD: a study in using universal deep features and transfer learning for automated AMD analysis. *Comput Biol Med*. 2017;82:80–86.
27. Ting DSW, Cheung YL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*. 2017;318(22):2211–2223.
28. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. *Med Image Anal*. 2019;58:101552, doi:10.1016/j.media.2019.101552. Epub 2019 Aug 31. PMID: 31521965.
29. Baur C, Albarqouni S, Navab N. MelanoGANs: high resolution skin lesion synthesis with GANs. Available at: https://arxiv.org/abs/1804.04338. Accessed April 17, 2021.
30. Mahapatra D, Antony B, Sedai S, Garnavi R. Deformable medical image registration using generative adversarial networks. In: *Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2018:1449–1453.
31. Nie D, Trullo R, Lian J, et al. Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans Biomed Eng*. 2018;65(12):2720–2730.
32. Beers A, Brown J, Chang K, et al. High-resolution medical image synthesis using progressively grown generative adversarial networks. Available at: https://arxiv.org/abs/1805.03144. Accessed April 17, 2021.
33. Frid-Adar M, Klang E, Amitai M, Goldberger J, Greenspan H. Synthetic data augmentation using GAN for improved liver lesion classification. In: *Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2018.
34. Rubin M, Stein O, Turko NA, et al. TOP-GAN: stain-free cancer cell classification using deep learning with a small training set. *Med Image Anal*. 2019;57:176–185.
35. Zhao H, Li H, Maurer-Stroh S, Cheng L. Synthesizing retinal and neuronal images with generative adversarial nets. *Med Image Anal*. 2018;49:14–26.
36. Borji A. Pros and cons of GAN evaluation measures. Available at: https://arxiv.org/abs/1802.03446. Accessed April 17, 2021.
37. Zhang X, Karaman S, Chang SF. Detecting and simulating artifacts in GAN fake images. In: *Proceedings of the 2019 IEEE International Workshop on Information Forensics and Security (WIFS 2020)*: Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2020.
38. Narayanaswamy A, Sakata LM, He GM, et al. Diagnostic performance of anterior chamber angle measurements for detecting eyes with narrow angles: an anterior segment OCT study. *Arch Ophthalmol*. 2010;128(10):1321–1327.
39. Teresa A, Aresta G, Mendonça L, et al. DR|GRADUATE: uncertainty-aware deep learning-based diabetic retinopathy grading in eye fundus images. *Med Image Anal*. 2020;63:101715.
40. Wang N, Wu H, Fan Z. Primary angle closure glaucoma in Chinese and Western populations. *Chin Med J*. 2002;115(11):1706–1715.