Semi-supervised generative adversarial networks for closed-angle detection on anterior segment optical coherence tomography images: an empirical study with a small training dataset

Ce Zheng[1], Victor Koh[2], Fang Bian[3], Luo Li[4], Xiaolin Xie[4], Zilei Wang[5], Jianlong Yang[6], Paul Tec Kuan Chew[2], Mingzhi Zhang[4]

[1]Department of Ophthalmology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China, Shanghai, China; [2]Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore; [3]Department of Ophthalmology, Deyang People's Hospital, Deyang, China; [4]Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong, Shantou University Medical College, Shantou, China; [5]Shanghai Children's Hospital, Shanghai, China; [6]Ningbo Institute of Industrial Technology, Chinese Academy of Sciences, Ningbo, China

Contributions: (I) Conception and design: C Zheng, M Zhang; (II) Administrative support: PTK Chew, M Zhang; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: V Koh, F Bian, X Xie, Z Wang, L Li; (V) Data analysis and interpretation: F Bian, X Xie, L Li, J Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

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

Background: Semi-supervised learning algorithms can leverage an unlabeled dataset when labeling is limited or expensive to obtain. In the current study, we developed and evaluated a semi-supervised generative adversarial networks (GANs) model that detects closed-angle on anterior segment optical coherence tomography (AS-OCT) images using a small labeled dataset.

Methods: In this cross-sectional study, a semi-supervised GANs model was developed for automatic closed-angle detection training on a small labeled and large unsupervised training dataset collected from the Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong (JSIEC). The closed-angle was defined as iris-trabecular contact beyond the scleral spur in AS-OCT images. We further developed two supervised deep learning (DL) models training on the same supervised dataset and the whole dataset separately. The semi-supervised GANs model and supervised DL models' performance were compared on two independent testing datasets from JSIEC (515 images) and the Department of Ophthalmology (84 images), National University Health System, respectively. The diagnostic performance was assessed by evaluation matrices, including the accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC).

Results: For closed-angle detection using clinician grading of AS-OCT imaging as the reference standard, the semi-supervised GANs model showed comparable performance, with AUCs of 0.97 (95% CI, 0.96–0.99) and 0.98 (95% CI, 0.94–1.00), compared with the supervised DL model (using the whole dataset) [AUCs of 0.97 (95% CI, 0.96–0.99), and 0.97 (95% CI, 0.94–1.00)]. When training on the same small supervised dataset, the semi-supervised GANs achieved performance at least as well as, if not better than, the supervised DL model [AUCs of 0.90 (95% CI: 0.84–0.96), and 0.92 (95% CI, 0.86–0.97)].

Conclusions: The semi-supervised GANs method achieves diagnostic performance at least as good as a supervised DL model when trained on small labeled datasets. Further development of semi-supervised learning methods could be useful within clinical and research settings.

Trial registration number: ChiCTR2000037892.

Introduction

Primary angle-closure glaucoma (PACG) is a major cause of irreversible vision loss (1-3). A closed-angle is characterized by appositional contact between the peripheral iris and trabecular meshwork. Gonioscopy is the current gold standard for evaluating the anterior chamber angle (ACA) and detecting the closed angle (4). However, gonioscopy requires direct contact with the patient's eye and is highly dependent on the examiner's expertise. Anterior segment optical coherence tomography (AS-OCT) is a noncontact technique used for imaging and quantitative measurement of the ACA with excellent reproducibility (5,6). However, AS-OCT has limitations, preventing its use as a solitary ACA screening system. For example, to evaluate ACA, clinicians have to manually mark the scleral spur, although some researchers had proposed the algorithm to detect the scleral spur, and modern AS-OCT modality can image the scleral spur more easily (7-9). Thus, developing an automated system would help overcome these limitations and allow closed-angle screenings of large populations.

Deep learning (DL) algorithms, a class of machine learning technology, can identify highly complex patterns in large, well-labeled datasets (10). In medicine and ophthalmology, supervised learning is the primary form of DL for image classification, for example, for detecting glaucoma, age-related macular degeneration (AMD), diabetic retinopathy (DR), and retinopathy of prematurity (ROP) (11-14). More recently, Fu *et al.* proposed a supervised DL system to detect iridotrabecular contact (ITC) with high accuracy [an area under the receiver operating characteristic curve (AUC) of 0.96] training on 8,270 AS-OCT ACA images (15). Using a training dataset with 4,036 AS-OCT images, Xu *et al.* reported a supervised DL system for detecting gonioscopic closed-angle closure with an AUC of 0.933 (16). However, creating a sizeable well-labeled dataset requires many resources and may not be available in most clinical practices.

Semi-supervised learning (SSL) is an attractive approach for addressing the lack of large, well-labeled datasets (17).

In contrast with supervised learning algorithms, SSL algorithms can improve performance by leveraging an unlabeled dataset when labeling is limited or expensive to obtain (18). For instance, Odena *et al.* suggested a semi-supervised generative adversarial networks (GANs) architecture. They demonstrated that this method's performance was at least comparable to (or better than) that of a standalone DL model when trained with few labeled examples (using only 25, 50, 100, and 1,000 labeled images from the public MNIST dataset, which has 60,000 images in total) (19). In this study, we propose, for the first time, to the best of our knowledge, a semi-supervised GANs method for automatically detecting closed angles in AS-OCT images and compared this approach with a supervised DL method based on a small label dataset. We present the following article in accordance with the STARD reporting checklist (available at http://dx.doi.org/10.21037/atm-20-7436).

Methods

This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the institutional review board (IRB) of the Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong (JSIEC) (identifier, 2018RY029-E01), China and National University Hospital System (NHG 292015/00788), Singapore and informed consent was taken from all the patients.

Clinical assessment

The primary study population consisted of adult phakic subjects consecutively examined in a hospital-based study from JSIEC between September 2014 and July 2017. The study methodology and details of the study population have been described previously. In brief, all subjects underwent a standardized ophthalmic examination including best-corrected visual acuity, refraction, slit-lamp examination,

and gonioscopic anterior chamber angle evaluation by a fellowship-trained glaucoma specialist (Goldmann 2-mirror lens, Haag-Streit AG, Bern, Switzerland). The angle was graded using 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 (20). Imaging was performed in darkroom conditions with Casia ASOCT (SS-1000 with a scan speed of 30,000 A-scans per second, Tomey Corporation, Nagoya, Japan). Inclusion criteria for the study included phakic eyes that received gonioscopy and AS-OCT imaging. AS-OCT imaging and gonioscopy were performed in one day. Eyes with corneal abnormalities (edema, pterygium, and other degenerative changes) and previous laser or intraocular surgery (iridotomy, iridoplasty, cataract surgery, or trabeculectomy) were excluded. In AS-OCT images, the closed-angle was defined as iris-trabecular contact beyond the scleral spur (21).

Datasets

AS-OCT images from 1,047 subjects were retrospectively collected from JSIEC. The raw dataset contains 17,845 Casia AS-OCT images. All AS-OCT images were reviewed and labeled by three fellowship-trained glaucoma specialists with more than ten years of experience blinded to all patient information. We excluded 1,029 images (64 eyes from 43 subjects) due to poor-quality images. The AS-OCT images were then split into two ACA images and resized to 128×128 pixels to reduce hardware demands during DL model training. After image processing, the AS-OCT dataset contained 33,864 Casia ACA images. After image grading and preprocessing, we further built three training datasets for semi-supervised GANs and DL models development: (I) a fully supervised dataset (33,864 ACA images with labels); (II) a small supervised dataset (20, 50, 100, 200, 400 and 1,000 ACA images randomly selected from the whole dataset); and (III) an unsupervised dataset (32,864 ACA images without labels) (*Figure 1*).

We also collected two independent testing datasets to evaluate the classification and generalization of the semi-supervised GANs model's performance using the same inclusion and exclusion criteria. In the testing datasets, we only included Casia AS-OCT images with a horizontal scan. The first testing dataset (hereinafter referred to as the "JSIEC testing dataset") was collected from an ongoing clinical trial (Chinese clinical trial registration number:

ChiCTR2000037892) in JSIEC between March 2019 and June 2020. The JSIEC testing dataset included 264 open-angle and 251 angle-closure ACA images. The second testing dataset, including 44 open-angle and 40 angle-closure ACA images, was enrolled from the National University Health System (hereinafter referred to as the "NUHS testing dataset") between January 2020 and March 2020. We used a customized software [Anterior Segment Analysis Program (ASAP)] to measure AS-OCT parameters between the JSIEC and NUHS testing datasets (22). The parameters included: AOD 750 (the length of the line segment between the cornea and iris at a 750 μm distance from the scleral spur), ACD (the anterior chamar depth), and TISA750 (the trabecular-iris space area at 750 μm anterior to the scleral spur).

Semi-supervised generative adversarial networks (GANs)

The details of the semi-supervised GANs architectures have been previously described by Odena *et al.* and Diaz-Pinto *et al.* (19,23). The standard GANs utilize two networks, a discriminator and a generator, involved in a minimax game to find the Nash equilibrium of these two networks (24).

$$\min_G \max_D V(G,D) = \\ \mathbb{E}_{x \sim p_{data}(x)}\left[\log D(x)\right] + \mathbb{E}_{z \sim p_z(z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad [1]$$

In Eq. [1], Ex~pdata(x) is the expectation over the training data (maximize log(D(x)) and Ez~pz(z) is the expectation over the data produced by the generator (minimize log(1 –D(G(z))).

Semi-supervised GANs are an extension of the GANs architecture. Instead of binary classification (real *vs.* fake in the standard GANs), the discriminator is transformed into a K+1-class (K = 2 here) classifier in semi-supervised GANs (19,25).

$$L_{semi-supervised} = L_{supervised} + L_{unsupervised} \quad [2]$$

In Eq. [2], $L_{sem-supervised}$ represents the semi-supervised loss function, and the total loss is used in optimization.

$$L_{supervised} = \mathbb{E}_{x,y \sim p_{data}(x,y)} \log\left(p_{model}\left(y|x, y < K+1\right)\right) \quad [3]$$

$$L_{unsupervised} = \\ -\left\{\mathbb{E}_{x \sim p_{data}(x)} \log D(x) + \mathbb{E}_{z \sim p_z(z)}\log\left(1 - D\left(G(z)\right)\right)\right\} \quad [4]$$

In Eq. [3], $L_{supervised}$ represents the supervised loss defined by the cross-entropy loss function as in a supervised learning
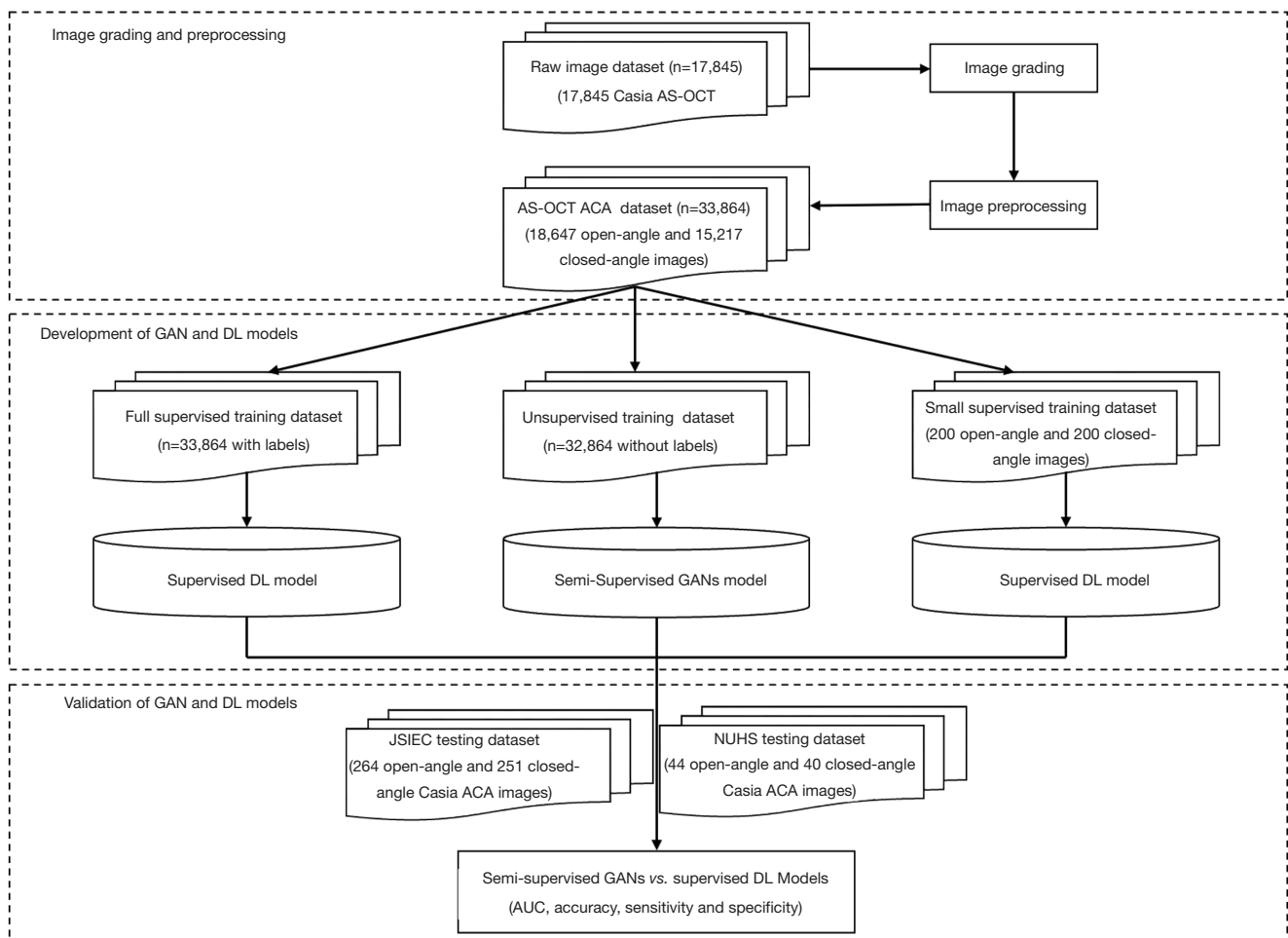
Page 4 of 11

Zheng et al. Semi-supervised GANs for closed-angle detection on AS-OCT

Figure 1 Workflow diagram showing the development of semi-supervised GANs and a supervised DL model to detect closed angles in AS-OCT images.

setting with K classes (open-angle *vs.* closed-angle in the current study). The unsupervised loss function ($L_{unsupervised}$) is the loss function of the standard GANs. As a result, this architecture allows unlabeled real data to contribute to learning, reducing the amount of labeling effort required to achieve a certain accuracy level.

A schema of the semi-supervised GANs architecture is shown in *Figure 2*. We followed the guidelines to construct the generator and discriminator described by Brownlee *et al.* (26). In this architecture, the generator is fed a 100x1 input vector (noise). There are then four transpose convolutional layers (kernel size =5, stride =2) with ReLU activation to scale to the appropriate 128×128 image size. The discriminator network is a similar network with a series of convolution layers (kernel size =5, stride =2). For the last output layer of the discriminator, we use one neuron for unsupervised classification (synthetic *vs.* real) and two neurons for supervised classification (open-angle *vs.* closed-angle).

Evaluating the discriminator of semi-supervised GANs for closed-angle detection

To evaluate the semi-supervised GANs model's performance for closed-angle detection, we trained two supervised DL models: the 1st supervised DL model (DL_Model_F) using the fully supervised training dataset and the 2nd supervised DL model (DL_Model_S) using the small supervised dataset. Kermany *et al.* described the details of the supervised DL model with the transfer learning technique (27). In brief, a Google Inception V3 architecture was adopted with weights pre-trained on the
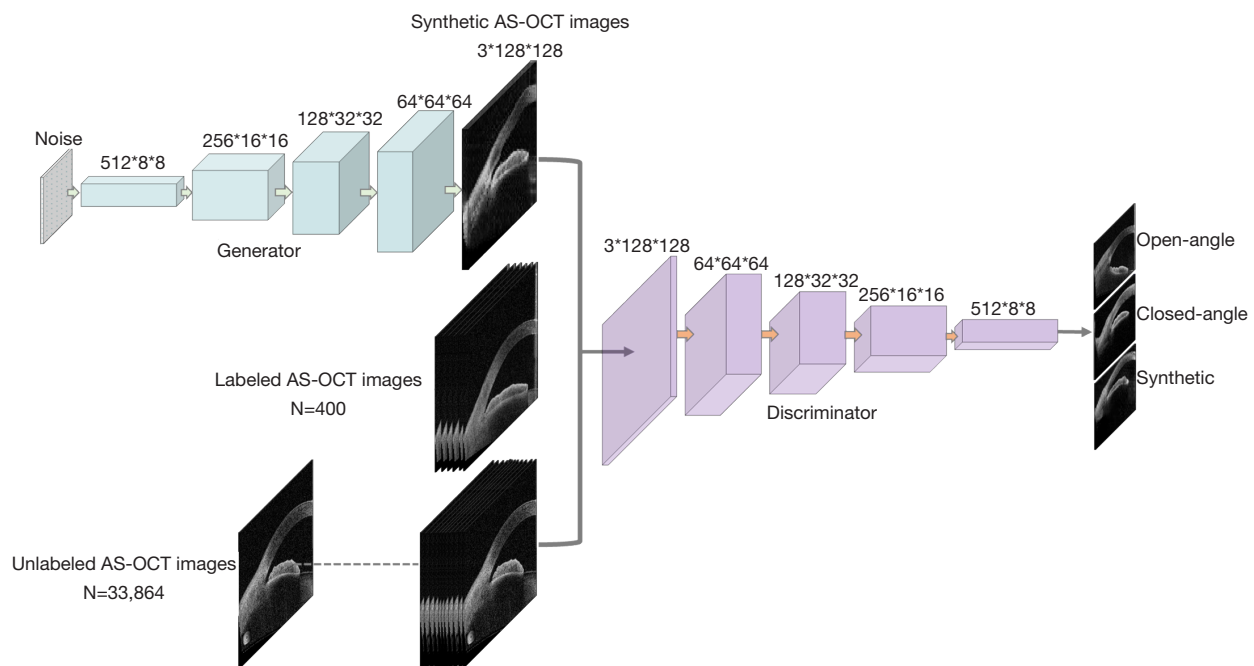
Figure 2 Schema of the semi-supervised GANs architecture.

ImageNet dataset (28,29). We further fine-tune the pre-trained model to recognize our specific classes by adding new neural network layers on the top. The new layers include one global average pooling layer, one hidden fully connected layer (including 256 neurons), and a final sigmoid classification layer to output. The small supervised training dataset was augmented through horizontal flipping, random rotation between 0 and 10 degrees, and random adjustments to saturation. We then assessed the discriminator's performance in the semi-supervised GANs and the supervised DL models in both testing datasets.

Despite promise during GANs training, model collapse is a failure case for GANs where the generator generates a limited diversity of images or images too similar to the original training images (30). To qualitatively evaluate synthetic and real images, we followed the guidelines to build t-distributed stochastic neighbor embedding (t-SNE) plots suggested by Van Der Maaten *et al.* (31). A DL network first extracts the high-dimensional features (100 features in the current study) from synthetic and real images. The t-SNE technique further reduces the dimensions of the features from 100 to 2.

We implemented the semi-supervised GANs and supervised DL models with Keras API (version 2.2.4) and the TensorFlow framework (Google, version 2.1.0) as the backend. The computer platform was equipped with an NVIDIA GTX 1080Ti 12 GB GPU, 128 GB RAM, and Intel Core i7-2700K 4.6 GHz CPU.

Statistical analysis

The evaluation matrices used to assess the performance of the semi-supervised GANs and the supervised DL models included accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC) with 95% confidence intervals (95% CIs). The accuracy, sensitivity, and specificity of the DL algorithms for detecting closed angles were computed according to the reference standard. All statistical analyses were performed using the Python (version 3.7) and Scikit_learn modules (Anaconda, version 1.9.12, Continuum Analytics)

Results

A total of 33,864 ACA images from 842 participants (mean age 60.4 years, 57.3% female) were included in the current study. We tested interobserver variability for closed-angle detection in randomly selected datasets (400 ACA images in total) and obtained kappa agreement better than 0.8. The JSIEC testing dataset consisted of 515 ACA images from

Table 1 AS-OCT biometric parameters between the supervised training dataset and the JSIEC and NUHS testing datasets

	JSIEC testing dataset			NUHS testing dataset		
	Open-angle (n=264)	Closed-angle (n=251)	P	Open-angle (n=44)	Closed-angle (n=40)	P
AOD750 (mm)*	0.47±0.22	0.10±0.10	<0.001	0.43±0.21	0.11±0.09	<0.001
ACD (mm)*	2.49±0.48	1.62±0.37	<0.001	2.36±0.54	1.87±0.32	<0.001
TISA750 (mm^2)*	0.28±0.13	0.02±0.02	<0.001	0.21±0.11	0.03±0.03	<0.001

*AOD750 (mm) was defined as the line segment's length between the cornea and iris at a 750 μm distance from the scleral spur. TISA750 was defined as a trabecular-iris space area 750 μm anterior to the scleral spur. ACD was defined as the anterior chamar depth.

Table 2 The diagnostic matrices of the semi-supervised GANs model for closed-angle detection with different training samples

	AUCs (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
20 samples	0.72 (0.68 to 0.76)	0.65 (0.61 to 0.69)	0.61 (0.57 to 0.65)	0.69 (0.65 to 0.73)
50 samples	0.79 (0.76 to 0.83)	0.72 (0.68 to 0.76)	0.65 (0.61 to 0.69)	0.78 (0.74 to 0.82)
100 samples	0.88 (0.86 to 0.91)	0.81 (0.74 to 0.82)	0.83 (0.81 to 0.87)	0.78 (0.74 to 0.82)
200 samples	0.93 (0.91 to 0.95)	0.86 (0.83 to 0.89)	0.84 (0.81 to 0.88)	0.88 (0.85 to 0.91)
400 samples	0.96 (0.94 to 0.98)	0.90 (0.87 to 0.93)	0.88 (0.85 to 0.91)	0.92 (0.90 to 0.94)
1,000 samples	0.97 (0.96 to 0.98)	0.91 (0.91 to 0.95)	0.90 (0.87 to 0.93)	0.92 (0.90 to 0.94)

AUCs, the area under the receiver operating characteristic curve; CI, confidence interval.

103 participants (mean age 63.8 years, 60.2% female). The NUHS testing dataset included 84 ACA images from 42 participants (mean age 61.4 years, 68.2% female). *Table 1* shows the AS-OCT biometric parameters between the JSIEC testing dataset and the NUHS training dataset.

For closed-angle detection using the clinician's grading of AS-OCT imaging as the reference standard, we first evaluated the semi-supervised GANs model's performance with decreasing supervised training samples of 1,000, 400, 200, 100, 50, and 20. The results reveal that the diagnostic capability of semi-supervised GANs underwent a performance degradation as the number of supervised datasets decreased (*Table 2*). Specifically, when trained with a large supervised dataset [n=1,000, accuracy =0.93 (95% CI, 0.91–0.95)], the semi-supervised GANs method had only slight improvements over the same method trained with smaller dataset [n=400, accuracy =0.92 (95% CI, 0.90–0.94)]. Considering similar results, we only used the semi-supervised GANs trained with 400 supervised samples for further experiments.

We then compared the diagnostic performance between the semi-supervised GANs and two supervised DL models (DL_Model_S and DL_Model_F). As shown in *Table 2* and *Figure 3*, the semi-supervised GANS model has performance close to the DL_model_F (training on the whole dataset). The semi-supervised GANs achieved better performance than the supervised DL_model_S when training on the same small supervised training dataset. In the two testing datasets, the semi-supervised GANs achieved AUCs of 0.97 (95% CI, 0.96–0.99) and 0.98 (95% CI, 0.94–1.00), which were higher than those of the supervised DL_model_S, with AUCs of 0.92 (95% CI, 0.84–0.96) and 0.90 (95% CI, 0.84–0.97), respectively (*Table 3* and *Figure 3*).

Regarding the qualitative evaluation, we used t-SNE to illustrate the embedding of high-dimensional data into a two-dimensional space (*Figure 4*). Each point represents a real or synthetic AS-OCT image projected from the 100-dimensional output of the CNN's last hidden layer into two dimensions. Two roughly separated point clouds could be observed in the t-SNE plot, which indicates that GANs can generate a continuum of novel synthetic AS-OCT images.

Discussion

In this study, we developed and tested a semi-supervised GANs model that detects closed angles on AS-OCT images. The semi-supervised GANs architectures contain three neural networks: (I) a supervised discriminator,
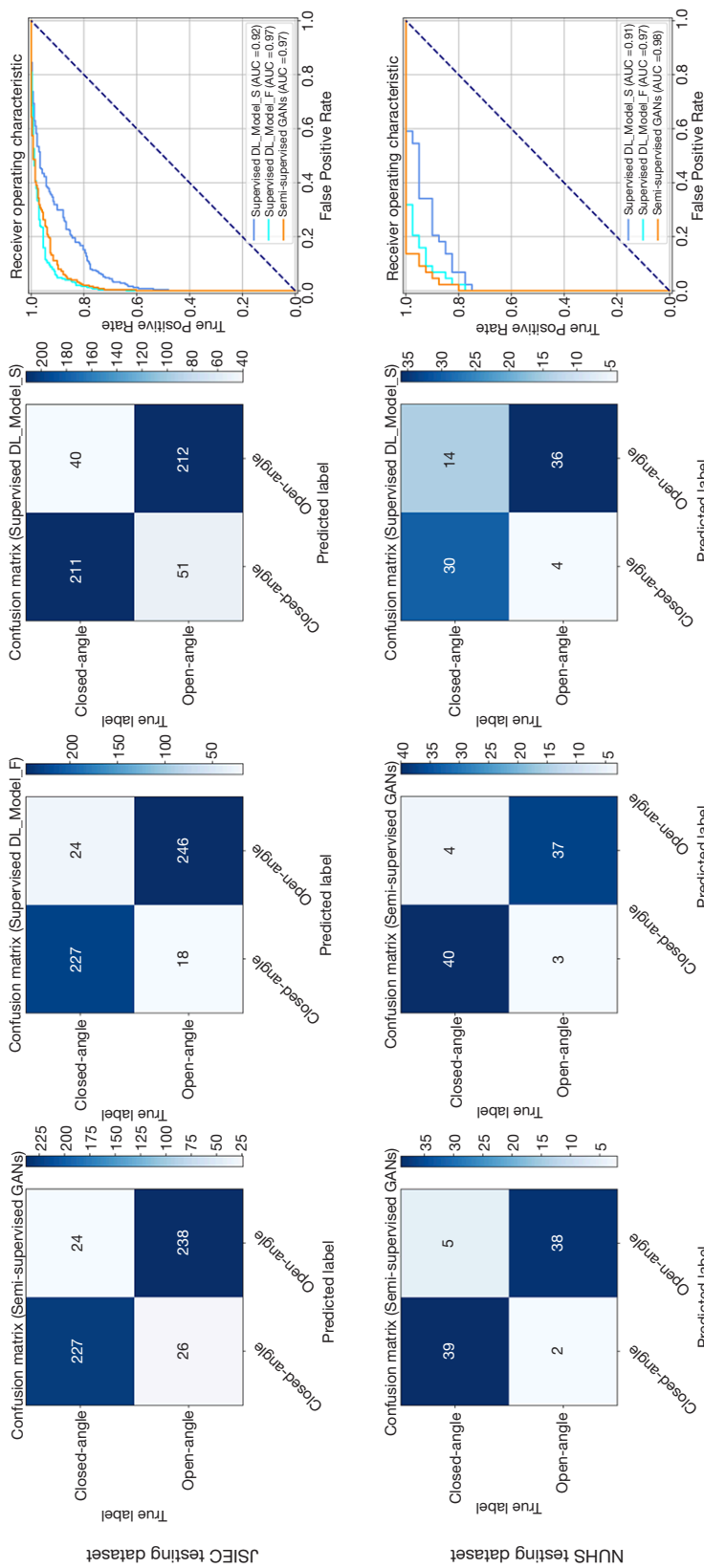
Figure 3 Confusion matrices and AUCs for semi-supervised GANs and supervised DL model in JSIEC and NUHS testing datasets. AUCs, the area under the receiver operating characteristic curve; GANs, generative adversarial networks; DL, deep learning.

Page 8 of 11

Zheng et al. Semi-supervised GANs for closed-angle detection on AS-OCT

Table 3 The diagnostic performance of semi-supervised GANs and two supervised DL models testing in the JSIEC and NUHS Datasets

	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
A: Testing in JSIEC Dataset			
Semi-supervised GANs	0.90 (0.87 to 0.93)	0.90 (0.87 to 0.93)	0.90 (0.87 to 0.93)
Supervised DL_Model_F	0.92 (0.90 to 0.95)	0.92 (0.87 to 0.93)	0.91 (0.89 to 0.94)
Supervised DL_Model_S	0.82 (0.79 to 0.85)	0.84 (0.81 to 0.87)	0.80 (0.77 to 0.84)
B: Testing in NUHS Testing Dataset			
Semi-supervised GANs	0.92 (0.86 to 0.98)	0.89 (0.82 to 0.96)	0.95 (0.90 to 0.99)
Supervised DL_Model_F	0.90 (0.84 to 0.96)	0.91 (0.85 to 0.97)	0.90 (0.84 to 0.96)
Supervised DL_Model_S	0.79 (0.70 to 0.88)	0.68 (0.58 to 0.78)	0.90 (0.84 to 0.96)

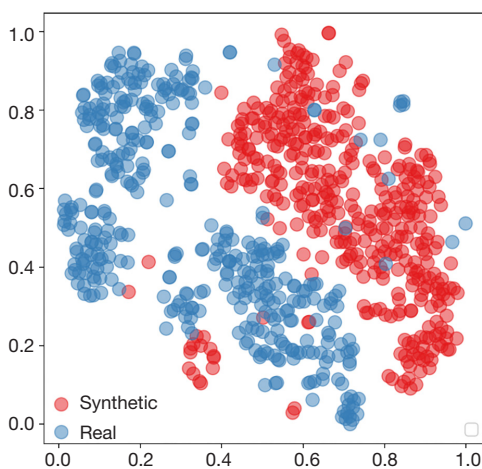GANs, generative adversarial networks; DL, deep learning.

Figure 4 High-dimensional features on 2D subspace by the t-SNE plot. Red and blue dots represent features of real and synthetic AS-OCT images, respectively.

(II) an unsupervised discriminator, and (III) a generator model (19,25). During training, our model simultaneously trains the discriminator for both the unsupervised GANs task using a large unlabeled dataset and the supervised classification task using a small labeled dataset. In two independent testing datasets, the semi-supervised GANs model's performance was acceptable when training with a relatively small dataset (400 images). Hence, our results suggest that a semi-supervised GANs method may perform at least as well as, if not better than, a supervised DL model in detecting closed angles on AS-OCT images when labeled data are scarce.

Although DL has improved the performance of many

challenging tasks in medical and ophthalmological imaging, there remain challenges and limitations to using DL. The DL models often benefit from being trained on large datasets. For example, Li *et al.* reported that a DL system could detect referable glaucomatous optic neuropathy based on 48,116 fundus photographs (32). Gulshan *et al.* collected datasets from two different countries for training a transfer learning DL model with an AUC of 0.974 for detecting diabetic retinopathy (33). When data are shared between different centers or counties, researchers need to consider IRB regulations and state privacy rules (34). Second, DL methods require high-quality labeled clinical data, which is challenging and requires expertise (35). Moreover, both qualitative image labeling and quantitative hand-crafted feature measurement can suffer from interobserver or intraobserver variability.

Semi-supervised learning is a kind of machine learning and occupies the middle ground between supervised learning (all training data are labeled) and unsupervised learning (all training data are unlabeled) (17). Given the low number of experts' labeled images and large unlabeled images in routine clinical settings, this technique can significantly help develop automatic assessment systems using medical images. Different research groups have proposed several SSL algorithms. Inés *et al.* presented an ML method that combines transfer learning with a new semi-supervised learning procedure (36). Their SSL method improves the accuracy of models up to 10% when working with partially annotated datasets. GANs are an artificial neural network architecture that was initially developed in the context of unsupervised learning (24). Recently, Odena *et al.* further suggested an SSL approach that involves

GANs by forcing the discriminator network to output class labels (19). Lecouat *et al.* adopted semi-supervised GANs for retinal disease diagnosis, and their results outperformed the supervised baselines by up to 15% when less than 30% of the training dataset was labeled (37). For cardiac abnormality classification in chest X-rays, Madani *et al.* also demonstrated that fewer data are required with semi-supervised GANs than with the conventional supervised DL model (38). In contrast to a previous DL study on AS-OCT images, we randomly selected a small labeled dataset (400 images in total) as the supervised training dataset, leaving other unlabeled ACA images into the unsupervised training dataset (16,39). Although our study cannot be directly compared to previous studies due to different reference standards of closed angles, our results still showed the potential application of this technique with excellent diagnostic performance.

It should be mentioned that the DL community has already used techniques, such as transfer learning and data augmentation, to address the limitations of small datasets (40). Transfer learning is a set of techniques that reuse a pre-trained model developed for other applications (e.g., training from ImageNet images) and further fine-tune the pre-trained model for a different target domain (e.g., detection of closed angles in the current study). When working with a limited number of images, transfer learning is usually combined with data augmentation, which is a technique that generates new training samples from the original dataset by applying transformations such as horizontal flipping, modest blurring, sharpening, and changes to saturation, brightness, contrast, and color balance. However, the above two techniques cannot leverage the information in both the labeled and unlabeled data. It is not surprising to find that, in the current study, the semi-supervised GANs achieved better performance than the supervised DL model using both transfer learning and data augmentation techniques.

The present study has several limitations. First, our semi-supervised GANs architecture can only synthesize images with 128×128 pixels lower than the Casia AS-OCT images' resolution. Our previous study reported a progressively grown GANs architecture to generate realistic OCT images with higher resolutions (e.g., 256×256 or above) (41). Second, we only used two independent testing datasets with small sample sizes, making the little difference in model performance challenging to interpret. Third, Fu *et al.* reported a supervised DL system to detect angle closure with high accuracy using ITC of 1/3 the width

of the trabecular meshwork (15). As different reference standards were used, our study cannot be directly comparable to Benjamin and Fu's studies. Additionally, gonioscopy was not the reference standard in this study. Future work is warranted to develop semi-supervised GANs for detecting eyes with gonioscopic angle closure.

In the current study, we developed a semi-supervised GANs method for detecting closed angles on AS-OCT images. The semi-supervised GANs method achieves diagnostic performance at least as good as, if not better than, the supervised DL model when trained on small labeled datasets. We hope this study prompts further development of SSL methods that could be best utilized within clinical and research settings.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-7436

Data Sharing Statement: Available at http://dx.doi.org/10.21037/atm-20-7436

Peer Review File: Available at http://dx.doi.org/10.21037/atm-20-7436

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-7436). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all

aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the institutional review board (IRB) of the Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong (JSIEC) (identifier, 2018RY029-E01), China and National University Hospital System (NHG 292015/00788), Singapore and informed consent was taken from all the patients.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. Ophthalmology 2014;121:2081-90

2. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol 2006;90:262-7.

3. Liang Y, Friedman DS, Zhou Q, et al. Prevalence and characteristics of primary angle-closure diseases in a rural adult Chinese population: The Handan eye study. Invest Ophthalmol Vis Sci 2011;52:8672-9.

4. Quek DT, Nongpiur M, Perera S, et al. Angle imaging: advances and challenges. Indian J Ophthalmol 2011;59 Suppl:S69-75.

5. Ang M, Baskaran M, Werkmeister RM, et al. Anterior segment optical coherence tomography. Prog Retin Eye Res 2018;66:132-56.

6. Tan AN, Sauren LDC, de Brabander J, et al. Reproducibility of anterior chamber angle measurements with anterior segment optical coherence tomography. Invest Ophthalmol Vis Sci 2011;52:2095-9.

7. Xu BY, Chiang M, Pardeshi AA, et al. Deep neural network for scleral spur detection in anterior segment OCT images: The Chinese American eye study. Transl Vis Sci Technol 2020;9:18.

8. Sakata LM, Lavanya R, Friedman DS, et al. Assessment of the scleral spur in anterior segment optical coherence tomography images. Arch Ophthalmol 2008;126:181-5.

9. Wong HT, Lim MC, Sakata LM, et al. High-definition optical coherence tomography imaging of the iridocorneal angle of the eye. Arch Ophthalmol 2009;127:256-60.

10. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.

11. Zheng C, Xie X, Huang L, et al. Detecting glaucoma based on spectral domain optical coherence tomography imaging of peripapillary retinal nerve fiber layer: a comparison study between hand-crafted features and deep learning model. Graefes Arch Clin Exp Ophthalmol 2020;258:577-85.

12. Burlina PM, Joshi N, Pekala M, et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. JAMA Ophthalmol 2017;135:1170-6.

13. Ting DSW, Cheung CYL, 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;318:2211-23.

14. 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;136:803-10.

15. Fu H, Baskaran M, Xu Y, 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.

16. Xu BY, Chiang M, Chaudhary S, et al. Deep Learning Classifiers for Automated Detection of Gonioscopic Angle Closure Based on Anterior Segment OCT Images. Am J Ophthalmol 2019;208:273-80.

17. Chapelle, O., Scholkopf, B., and Zien A. Semi-supervised learning. MIT Press; 2006.

18. Oliver A, Odena A, Raffel C, et al. Realistic evaluation of deep semi-supervised learning algorithms. In: Advances in Neural Information Processing Systems [Internet]. Neural information processing systems foundation; 2018. p. 3235–46. Available online: https://github.com/brain-research/realistic-ssl-evaluation

19. Odena A. Semi-Supervised Learning with Generative Adversarial Networks. 2016;1–3. Available online: http://arxiv.org/abs/1606.01583

20. Foster PJ, Buhrmann R, Quigley HA, Johnson GJ. The

definition and classification of glaucoma in prevalence surveys. Br J Ophthalmol 2002;86:238-42.

21. Baskaran M, Iyer J V., Narayanaswamy AK, et al. Anterior segment imaging predicts incident gonioscopic angle closure. Ophthalmology 2015;122:2380-4.

22. 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 Vis Sci 2012;53:6756-62.

23. Diaz-Pinto A, Colomer A, Naranjo V, et al. Retinal Image Synthesis and Semi-Supervised Learning for Glaucoma Assessment. IEEE Trans Med Imaging 2019;38:2211-8.

24. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Advances in Neural Information Processing Systems 2014. p. 2672–80.

25. Salimans T, Goodfellow I, Zaremba W, et al. Improved techniques for training GANs. In: Advances in Neural Information Processing Systems 2016. p. 2234–42.

26. Brownlee J. How to Implement a Semi-Supervised GAN (SGAN) From Scratch in Keras, Machine Learning Mastery [Internet]. 2019 [cited 2020 May 26]. Available online: https://machinelearningmastery.com/semi-supervised-generative-adversarial-network/

27. Kermany DS, Goldbaum M, Cai W, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 2018;172:1122-1131.e9.

28. Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016. p. 2818–26.

29. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings 2015.

30. Mao Q, Lee HY, Tseng HY, Ma S, et al. Mode seeking generative adversarial networks for diverse image synthesis. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society 2019. p. 1429–37.

31. Van Der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;9:2579-605.

32. Li Z, He Y, Keel S, et al. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. Ophthalmology 2018;125:1199-206.

33. Gulshan V, Peng L, Coram M, et al. Novel Low-Shot Deep Learning Approach for Retinal Image Classification With FewExamples. JAMA 2016;316:2402-10.

34. McMillan T. CCNA® Security Study Guide. CCNA® Security Study Guide 2018.

35. Ting DSW, Liu Y, Burlina P, et al. AI for medical imaging goes deep. Nat Med 2018;24:539-40.

36. Inés A, Domínguez C, Heras J, et al. Biomedical image classification made easier thanks to transfer and semi-supervised learning. Comput Methods Programs Biomed 2021;198:105782.

37. Lecouat B, Chang K, Foo C-S, et al. Semi-Supervised Deep Learning for Abnormality Classification in Retinal Images 2018;1–9. Available online: http://arxiv.org/abs/1812.07832

38. Madani A, Moradi M, Karargyris A, et al. Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation. Proc Int Symp Biomed Imaging 2018;2018:1038-42.

39. Fu H, Xu Y, Lin S, et al. Angle-Closure Detection in Anterior Segment OCT Based on Multilevel Deep Network. IEEE Trans Cybern 2020;50:3358-66.

40. Razavian AS, Azizpour H, Sullivan J, et al. CNN features off-the-shelf: An astounding baseline for recognition. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops 2014.

41. Zheng C, Xie X, Zhou K, et al. Assessment of generative adversarial networks model for synthetic optical coherence tomography images of retinal disorders. Transl Vis Sci Technol 2020;9:29.