A GAN-based Domain Adaptation Method for Glaucoma Diagnosis

Yunzhe Sun[1] , Gang Yang* [1], Dayong Ding[2], Gangwei Cheng[3], Jieping Xu[1], Xirong Li[1]
[1]AI & Media Lab, School of Information, Renmin University of China
[2]Vistel AI Lab, Visionary Intelligence Ltd.
[3]Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College.
* Corresponding Author, Email: yanggang@ruc.edu.cn

***Abstract*—Domain adaptation is an important research topic in the field of computer vision, where the goal is to solve the difference of data distribution between different scenarios of the same task. In recent times, adversarial learning method becomes a mainstream approach to generate complicated images across diverse domains through optimizing deep networks, and it can also improve the recognition accuracy rate of deep networks despite existing domain shift or dataset bias. However, there are few effective efforts of domain adaptation for the disease diagnosis on fundus images. Fundus images are normally captured on different medical devices with different rules. When diagnosing glaucoma, there is a serious homogeneous domain shift, which means feature spaces between target domain and source domain images have a distribution shift although they are very similar. We propose a unified framework to solve this problem. Previous studies have shown that glaucoma can be monitored by analyzing the optic disc/cup and its surroundings. So we exploit a novel reconstruction loss which not only leverages unsupervised data to bring the source and target distributions closer but also keeps original target domain images label unchanged. The experimental results on several public and private datasets demonstrate that our method could increase the classification accuracy of glaucoma diagnosis.**

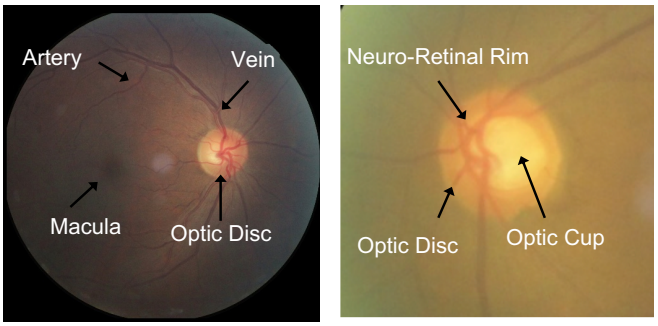
***Index Terms*—domain adaptation, glaucoma diagnosis, image synthesis, DAGD**

I. Introduction

When trained on large-scale datasets, deep convolutional neural networks (DCNNs) can learn effective representations which are largely useful across a variety of tasks and visual domains [1]–[3]. However, due to a problem known as dataset bias or domain shift [4], classification models trained along with the representations on one dataset can not perform equally well on other datasets or other classfication tasks [1], [5]. The common solution is to further fine-tune these models on task-specific datasets. But it is generally difficult and expensive to obtain enough labeled data to properly fine-tune a large number of parameters employed by deep multi-layer networks.

Domain adaptation methods try to weaken the harmful influence of domain shift. Recent domain adaptation methods try to learn effective transformations that map both source domain and target domain into a common latent feature space. Normally, this is achieved by optimizing the representations to minimize some measure of domain shift such as the maximum mean discrepancy [6], [7] or correlation distances [8], [9]. A feasible solution is to reconstruct the target domain from source representation [10].

This work was supported by the Beijing Natural Science Foundation (No. 4192029), and the National Natural Science Foundation of China (61773385, 61672523).



(a) (b)

Fig. 1. Fundus images. (a) Main structure of fundus image and (b) main structure of optic disc region. An abnormal size of the cup with respect to the optic disc is a characteristic of a glaucomatous eye.

embodimentembodimentAdversarial adaptation method has become a progressively popular embodiment of this approach. This method attempts to minimize an approximate domain discrepancy distance through an adversarial objective with respect to a domain discriminator. Normally, adversarial adaptation is directly bound up with generative adversarial learning [11], which makes generator and discriminator play against each other. The generator synthesizes images in a way that deceives the discriminator, which in turn tries to distinguish them from real images. In the field of domain adaptation, this theory has been employed to ensure that the network can not distinguish between the distributions of its source and target domain data [9], [12], [13]. However, these algorithms make different design choices, such as whether to use a generator, or to share weights across domains and which loss function to choose. For example, [14] and [12] share weights and learn a symmetric mapping of both source and target images into a shared feature space, while [13] separates some layers to learn a partially asymmetric mapping.

Nevertheless, there are currently no suitable domain adaptation methods that adapt to the fundus images for glaucoma diagnosis. Generally, fundus images are often captured on different medical devices and the image-level style is basically

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

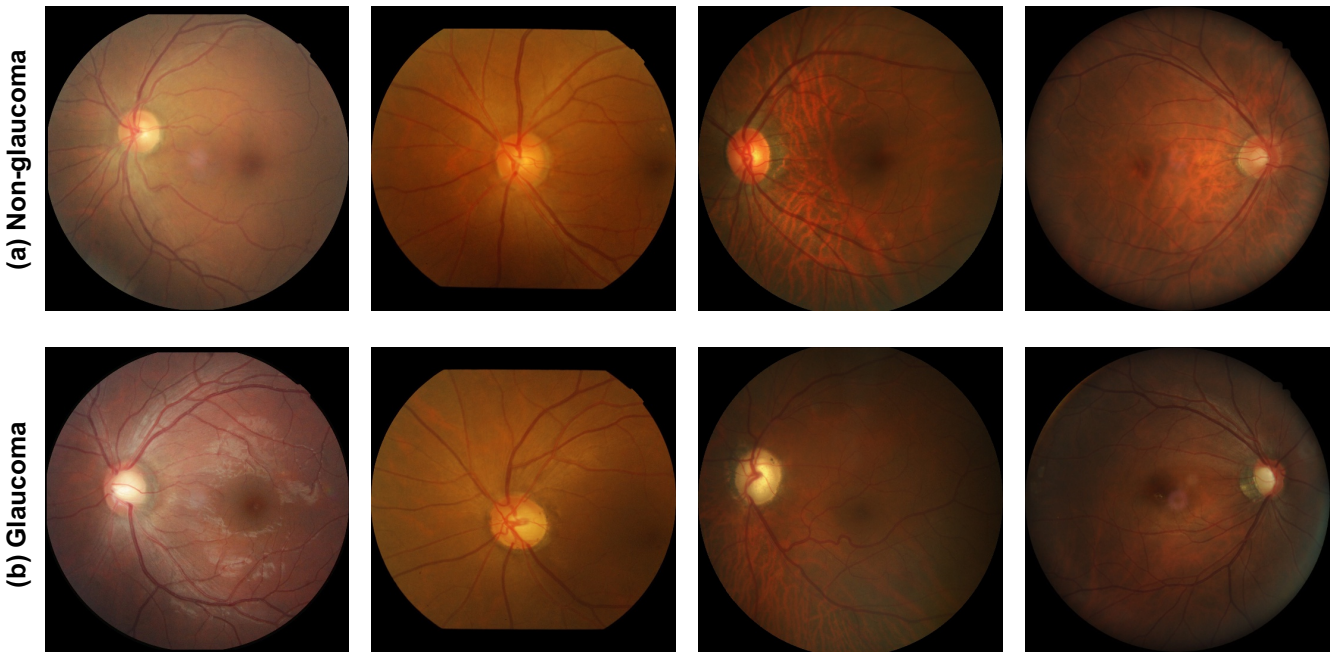


Fig. 2. Fundus images from four different datasets. (a) are non-glaucoma fundus images and (b) are glaucomatous fundus images. Each column of images comes from different datasets.

similar. For example, as shown in Fig. 2, comparing the fundus images in the first column and second column, the difference lies in the position of the optic disc and optic cup. And for the first and the third column, the most obvious difference lies in the overall brightness. Various studies have shown that glaucoma can be diagnosed by analyzing the optic disc/cup of fundus images, as shown in Fig. 1, so these differences are not the most important factor to diagnose glaucoma for an ophthalmologist. But for classification models that trained on different datasets, these differences have vital influences to make the classification.

In this paper, we propose an effective framework based on Generative Adversarial Networks (GAN), named DAGD, to solve this domain bias problem, which can learn the distribution of the features of source domain images and apply to target domain images. Besides, our framework exploits a novel GAN-loss to try to keep the original target image labels unchanged. The DAGD is mainly composed of two models. One is an unsupervised generative adversarial model which we called UDAGAN, and the other is a classification model pre-trained with source images. UDAGAN is responsible for transforming the target domain images into source-like images. We employed a generator in UDAGAN to synthesize a source-like target domain image that has the same distribution with source domain images and keep the original label unchanged. And the classification model is only used to classify the source-like fundus image glaucomatous or not. The main contributions of this paper are summarized as follows:

- We propose a novel domain adaptation method (DAGD) with an unsupervised generative adversarial network (UDAGAN) which is effective and suitable for glaucoma diagnosis on fundus images in unsupervised domain adaptation.
- We design a new reconstruction loss function that can not only learn the characteristics of the source domain but also keep the original label unchanged, in glaucoma diagnosis of fundus images.
- To our knowledge, our work firstly provides an attempt to perform domain adaptation on fundus images. Results on multiple medical image datasets show that our attempts are effective.

II. Related Work

A. Domain Adaptation

There has been a wide range of previous work on domain transfer learning, e.g. [15]. Recent work focuses on transferring deep convolutional neural networks representations from labeled source domain dataset to target domain, in which labeled data is scarce or non-existent. For unlabeled target domain images, the main algorithm is to guide feature learning by minimizing the difference of data distributions between source and target domain [9], [10], [12], [14].

To achieve this goal, several methods have used the Maximum Mean Discrepancy (MMD) [15] loss. MMD calculates the norm of the difference between the means of source domain and target domain. In addition to the common classification loss of the source, the DDC method [6] uses MMD to learn a representation that is both domain and discriminative invariant. Deep Adaptation Network (DAN) [7] applies MMD to the layers embedded in the regenerative kernel Hilbert space, effectively matching the higher-order statistics of two distributions. On the contrary, depth correlated alignment

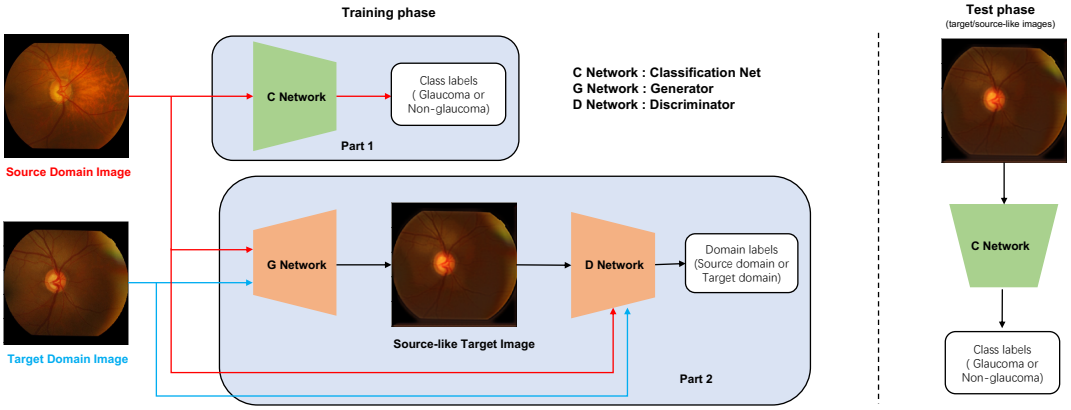


Fig. 3. The overview of the proposed framework DAGD and the inner model mainly consisting of two parallel parts. (1) In the training phase, Part 1 is the classification model training along with the source domain images. Part 2 is the UDAGAN model, which learns the source domain images style features and transforms target domain images to source-like images while keeping optic disc, optic cup and label unchanged. (2) In the test phase, we use a classification model to predict the label of input images. Note: our classification model is trained only using source images and labels. The part 2 UDAGAN does not need the classification model or any labels of source images and target images.

(CORAL) method [9] is proposed to match the mean and covariance of two distributions.

Other methods have chosen an adversarial loss to minimize domain shift, learning a representation that can distinguish between source labels but not domains. [12] proposes increasing a domain classifier (a fully connected layer) that predicts the binary domain label of the inputs and designs a domain confusion loss to encourage its prediction to be closer to a uniform distribution over binary labels. The gradient reversal algorithm (ReverseGrad) [14] proposes to treat domain invariance as a binary classification issue, but maximizes the loss of the domain classifier by reversing its gradients directly. However, these methods tend to focus on the problem of domain adaptation between two different data distributions, regardless of how much the original images will change. This is a serious problem for medical images, and important diagnostic evidence for medical images may be lost.

B. Generative Adversarial Networks

Generative Adversarial Networks (GANs) [11], [16] have been the most promising method for image generation [17]–[19], data augmentation [20] and domain adaptation [21], [22]. GANs are deep learning methods comprised of two networks. One is called generator and the other is called discriminator. During the training phase, the generator tries to synthesize realistic images while the discriminator tries to distinguish the generated from real data.

With the gratitude to GANs, image-level adaptation methods have been developed to domain shift at the input stage of DCNNs. Some methods first train a classification model on the source domain, and then transform the target domain image into source-like ones so that it can be tested using the pre-trained source-task model [23]–[25]. Instead, other methods try to transfer the source domain images into the appearance of target images [26]–[28]. Then the transformed target-like images with source label are used to train a target-task model which could perform well in the target domain. This has also been used in medical fundus image analysis [29]. With the amazing success of unpaired image-to-image transformation, e.g. CycleGAN [30], many previous image adaptation works are based on modified CycleGAN with applications in both domain datasets [23], [27].

At the same time, the feature-level adaptive methods have also been studied to reduce the domain offset by extracting domain invariant features in DCNNs. Pioneer works attempt to minimize the distance between domain statistics, such as the layer activation correlation [9] and maximum mean distance [31]. Later, the representative methods of DANN [32] and ADDA [22] achieve advanced feature adaptation through adversarial learning and discriminators to distinguish feature spaces across domains. The image and feature adaptations address domain shift from different perspectives to the DCNNs, which are in fact complementary to each other. Combining these two adaptive strategies to achieve a stronger domain adaption technique has made successful progress, e.g. [33].

Considering the homogeneous domain shift in fundus images, in addition to learning the data distribution of the source domain, we should keep the label of the original target domain images unchanged. To tackle this challenging adaptation for the diagnosis task, we propose to exploit a novel architecture and a reconstruction loss to constrain the changes during training.

III. DAGD

In this section, we introduce the unified framework DAGD and discuss our optimization method in detail. An overview of our framework is shown in Fig. 3. Our framework is mainly composed of two parallel parts. One is generative model UDAGAN (Part 2 in Fig. 3) and the other is the classification network (Part 1 in Fig. 3). The UDAGAN learns the source domain style features and transforms target domain images to source-like images with original image labels unchanged. The classification model is trained along with the source domain dataset. In DAGD, we assume that we have access

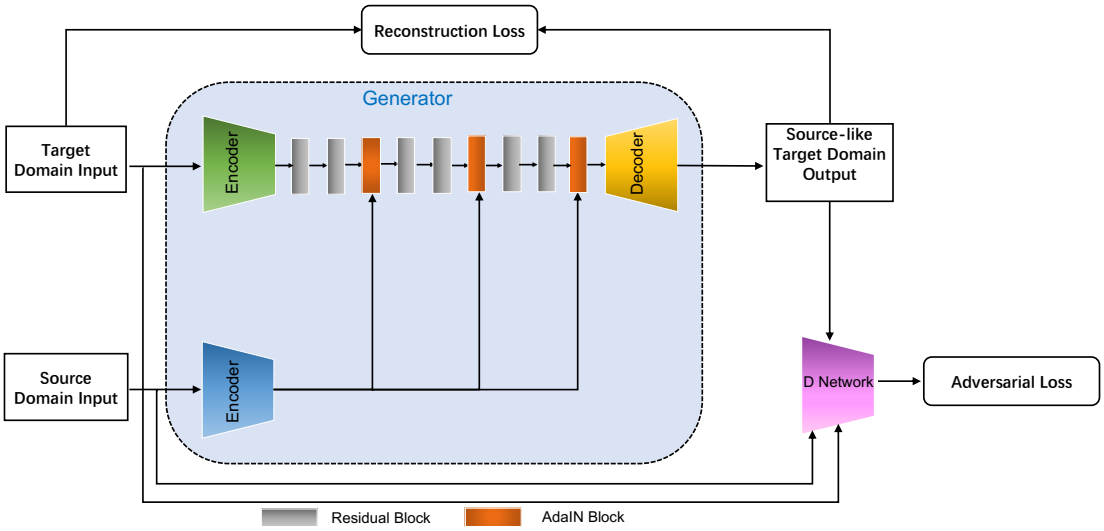


Fig. 4. The overview of UDAGAN. The generator encoders architecture and decoder architecture are the same as pix2pix. Encoders and decoder use modules of the form convolution-BatchNorm-ReLu.

to source images X_s drawn from a source domain distribution $p_s(x)$, as well as target images X_t drawn from target domain distribution $p_t(x)$. Our goal is to learn a mapping function $\mathcal{F}$ which can transform target domain images to source-like images and assure the main information steady for diagnosis as well as keep target images label unchanged without any supervised information. With this approach, we can directly use the source-task classification model to predict the labels of target domain images without fine-tuning the model.

A. UDAGAN for Domain Adaptation

An overview of our proposed UDAGAN architecture is shown in Fig. 4. The generator is mainly composed of two encoders and a decoder. The target encoder is responsible for learning latent features from target domain images, and the source encoder aims to extract style representations from source domain images. To combine the latent feature of target images with style representations extracted from source images, we adopt the Adaptative Instance Normalization method [34] and residual block used in [35]. We connect an AdaIn block after every two residual blocks. Finally, a decoder is responsible for generating realistic source-like images with representations from the last AdaIn block. According to the operation of AdaIN [34], we define our AdaIN block as follows:

$$AdaIN(x_t, x_s) = \sigma(x_s)(\frac{x_t - \mu(x_t)}{\sigma(x_t)}) + \mu(x_s) \qquad (1)$$

where x_t, x_s denotes the image sampled from target domain and source domain. $\mu(\cdot)$ and $\sigma(\cdot)$ denote mean and variance across spatial dimensions. Our generator can not only learn the context features of target samples but also extract the style information of source samples through AdaIN blocks.

For the discriminator, we use 70 x 70 PatchGANs [35]–[37] which aims to classify whether 70 x 70 overlapping image patches are from source domain or target domain. Such a patch-level discriminator architecture has fewer parameters than a full-image discriminator and can work on arbitrarily-size images in a fully convolutional fashion [35].

B. Model Losses

Like regular CGANs [38], the DAGD model emulates a competition, where the generator attempts to produce realistic source-like images, while the discriminator classifies between images from the source domain or the target domain. The main goal of the UDAGAN model is to maximize the miss-classification error of discriminator while the generator produces more realistic source-like images trying to fool the discriminator. This competition is also called a two-player minimax game and it can be described as follows:

$$\begin{aligned}\min_G \max_D V(G,D) = &\mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] \\ &+ \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]\end{aligned} \qquad (2)$$

where $\mathbb{E}_{x\sim p_{data}(x)}$ is the expectation over the real data and $\mathbb{E}_{z\sim p_z(z)}$ is the expectation over the data produced by generator. $D(x)$ represents the probability that x came from the real data rather than the data synthesized by the generator and $G(z)$ represents the probability of z produced by the generator. Therefore, the system is trained to minimize $\log(1 - D(G(z)))$ and maximize $\log(D(x))$.

However, the standard adversarial loss can not guarantee the target domain images label unchanged and keep diagnosis information steady. So we introduce a reconstruction loss to deal with this issue. And the final loss is defined as follows:

$$\min_G \max_D \mathcal{L}_{adv}(G,D) + \gamma \mathcal{L}_{rec}(G) \qquad (3)$$

where $\mathcal{L}_{adv}$ indicates the adversarial loss and we choose WGAN-GP loss [39] instead of traditional adversarial loss in Eq. (2), which we find useful to increase training stability. L_{rec} is a reconstruction loss and γ is a hyperparameter to balance the WGAN-GP loss [39] and the reconstruction loss.

a) Overview of WGAN-GP Loss: WGAN-GP [39] is proposed for overcoming the disadvantages of WGAN [40]. In the actual experimental process of WGAN, there are still problems with difficult training and slow convergence. So Gulrajani *et al*. [39] propose a solution, which is called the gradient penalty. Adding a gradient penalty term makes the weight in the middle of the original data and the generated data as small as possible, which is equivalent to converting the fixed threshold of WGAN into a variable threshold. This method makes the training of GANs more stable.

b) Reconstruction Loss: We want to ensure that our method does not change the target domain images label and the information for the diagnosis, which will lead to the loss of some medical characteristics. So we add a reconstruction loss to make a constraint. The reconstruction loss is defined as follows:

$$\mathcal{L}_{rec} = \|G(x_t, x_s) - x_t\|^2 \tag{4}$$

where $G(x_t, x_s)$ means the source-like image produced by generator, and x_t is the original target domain image. The loss function we choose is the Mean Squared Error.

IV. Experiments

In this section we describe experimental details of fundus images domain adaptation. We evaluate our method using four different public and private fundus datasets and we report the experiment results of each dataset. Furthermore, we use t-SNE [41] to visualize and analyze the distribution of source images, original images and source-like images generated by UDAGAN.

A. Datasets

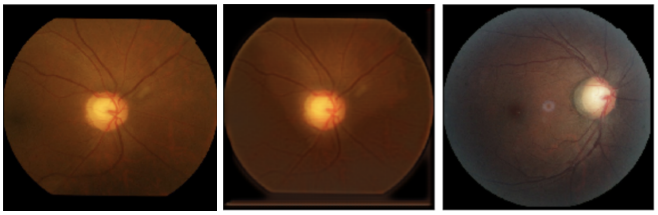
The experiments are conducted on four datasets: Drishti-GS dataset [42], REFUGE challenge dataset[1], PrivateDataset1 and PrivateDataset2. All these datasets are described in detail in Table I.

TABLE I
DATASETS USED TO TRAIN AND TEST IN OUR DAGD METHOD

Dataset	Glaucoma		Non-glaucoma		Total	
	train	test	train	test	train	test
Drishti-GS	32	38	18	13	50	51
REFUGE	40	40	360	360	400	400
PrivateDataset1	31	32	269	68	300	100
PrivateDataset2	958	131	7285	899	8243	1030

- **Drishti-GS:** The Drishti-GS [42] contains 101 images (training and test set). We use training set to train our UDAGAN and leverage test set to evaluate our method. The training set of Drishti-GS contains 32 glaucomatous images and 18 non-glaucoma images. The test set contains 38 glaucomatous images and 13 non-glaucoma images.
- **REFUGE challenge:** The REFUGE challenge dataset is composed of 1200 images(training, validate and test set). We choose the training set and validation set from REFUGE challenge. Both of the training and validation set are composed of 40 glaucomatous and 360 non-glaucoma images.
- **PrivateDataset1:** The PrivateDataset1 contains 400 fundus images in total. We divide 300 images of the dataset into a training set and the others into a test set. The training set contains 31 glaucomatous images and 269 non-glaucoma images. The test set contains 32 glaucomatous images and 68 non-glaucoma images.
- **PrivateDataset2:** PrivateDataset2 includes 9273 images, we split 8243 images into training set and 1030 images into test set. The images of the training set are not only used as source domain data to train UDAGAN model, but also used to train our classification networks.

[1]https://refuge.grand-challenge.org/Home/



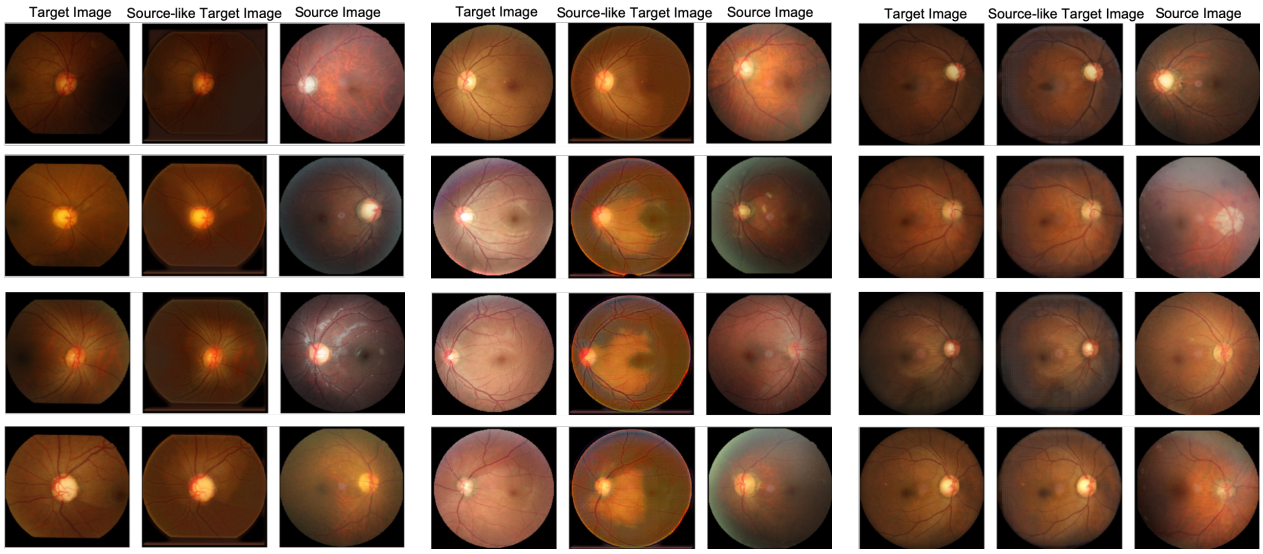
(a) Target (b) Synthesized (c) Source

Fig. 5. Target image, synthesized source-like image, and source image. (a) is sampled from Drishti-GS dataset. (b) is the synthesized image combining Drishti-GS features with style features of PrivateDataset2. (c) is sampled from PrivateDataset2 dataset.

B. Experiment Setup and Implementation

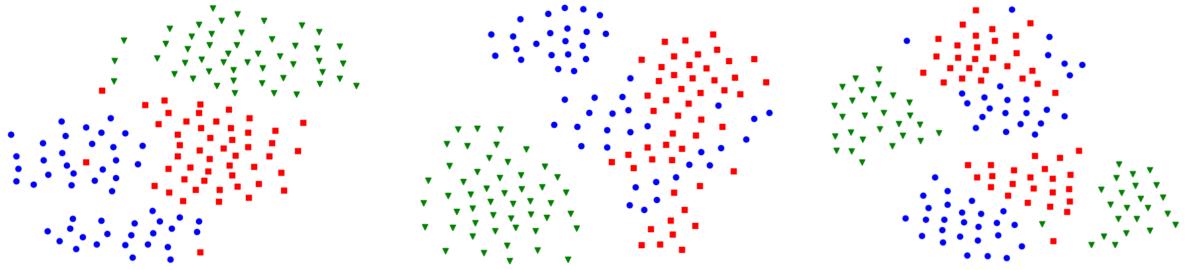
a) Classification Model: Our classification model is trained along with PrivateDataset2. The classification model architecture we use is ResNet-50 [43], which has been proved to be an effective architecture in the classification of glaucoma images [44]. In the procedure of data pre-processing, we resize PrivateDataset2 images to 512x512 pixel and flip the images horizontally with a probability of 0.5. And in the training phase, due to the imbalance of the proportion of positive and negative samples, we set a weighted random sampler to constrain the sampling process. The weight of the non-glaucoma fundus images is set to 1, and the weight of glaucomatous fundus images is set to 9. The optimization we use is the Adam solver [45] with $\beta_1 = 0.5$ and $\beta_2 = 0.999$ and batch size is 48. The network is trained with a fixed learning rate of 0.0001.

b) UDAGAN: Our generative model UDAGAN is composed of a generator and a discriminator. The generator encoders architecture and decoder architecture are the same as pix2pix [35]. In the intermediate stage, however, we add three AdaIn blocks [34] and connect an AdaIn block after every two residual blocks, so that the entire network can not only retain the information of the original target domain images but also learn the style of source domain images. For the discriminator, we used 70 x 70 PatchGANs [35]–[37], which has been proved to be an effective architecture in the field of image-to-image translation.



(a) Drishti-GS → PrivateDataset2 (b) REFUGE → PrivateDataset2 (c) PrivateDataset1 → PrivateDataset2

Fig. 6. Target images, synthesized source-like images, and source images.(a) The adaptation from Drishti-GS to PrivateDataset2, (b) the adaptation from REFUGE to PrivateDataset2 and (c) the adaptation from PrivateDataset1 to PrivateDataset2.



(a) Drishti-GS (b) REFUGE (c) PrivateDataset1

Fig. 7. t-SNE plots of the data distribution before and after domain adaptation in three datasets. (a) is Drishti-GS data distribution and (b) indicates REFUGE data. (c)is the PrivateDataset1. Green triangle indicates the original data distribution before domain adaptation and red square indicates the synthesized data distribution after domain adaptation. The blue dot denotes the PrivateDataset2 data distribution.

In all the experiments, we set $\gamma = 2$ in Eq. (3). We use the Adam solver [45] with $\beta_1 = 0.5$ and $\beta_2 = 0.999$. For data augmentation, we flip the images horizontally with a probability of 0.5 and resize the images to 512 x 512 pixel. The batch size is set to 4 for all experiments. We train our model with a fixed learning rate of 0.0001 until 1000 epochs and linearly decay up to 3000 epochs. We also use weight decay at a rate of 0.0001. The weight is initialized from a zero-centered non-glaucoma distribution with a standard deviation of 0.02.

For the adaptation from Drishti-GS to PrivateDataset2, due to the proportion of glaucomatous and non-glaucoma samples in Drishti-GS is closed to $2:1$, we choose the weight random sampler to set the weight of 1 for the glaucoma samples and 2 for non-glaucoma samples. For the adaptation from REFUGE to PrivateDataset2, the proportion of positive and negative samples in REFUGE is closed to $1:9$, so we set different weights between glaucoma images and non-glaucoma images, glaucoma sample weight is set to 9 and non-glaucoma is set to 1. The proportion of glaucoma and non-glaucoma in PrivateDataset1 is closed to $1:16$, so the glaucoma sample weight is set to 16 and non-glaucoma is set to 1.

C. Experimental Results

We experimentally validate our proposed method in an unsupervised domain adaptation task on three medical fundus datasets, Drishti-GS, REFUGE and PrivateDataset1. The PrivateDataset2 is used to train the classification networks and as the source domain to train UDAGAN. Images sampled from each dataset and the synthesized source-like target images are visualized in Fig. 5 and Fig. 6.

a) Visualization: From Fig. 6(a), the difference between the Drishti-GS and PrivateDataset2 is obvious and it is mainly reflected in the position of the optic disc and optic cup. Comparing the synthesized source-like images with original images (details are shown in Fig. 5), the optic disc and cup do not change a lot. The main changes are concentrated in the background which is not the determining factor for classification. And from Fig. 7 (a), we can find out the data

TABLE II
ACCURACY(%), SENSITIVITY(%) AND SPECIFICITY(%) VALUES ON THE THREE DATASETS. THE RESNET-SOURCE ONLY COLUMN REPRESENTS TARGET IMAGES PREDICTION RESULTS ON THE PRE-TRAINED SOURCE-TASK MODEL

Dataset	Assessment	ResNet-Source only	DAGD
Drishti-GS → PrivateDataset2	Accuracy	56.25	89.58
	Sensitivity	63.64	75.00
	Specificity	54.05	94.44
REFUGE → PrivateDataset2	Accuracy	88.75	89.75
	Sensitivity	70.00	72.50
	Specificity	90.83	91.66
PrivateDataset1 → PrivateDataset2	Accuracy	79.00	82.00
	Sensitivity	78.13	81.25
	Specificity	79.41	82.35

distribution of original Drishti-GS images (green triangle) has been transformed to a different distribution area (red square) which is closed to source domain data distribution (blue dot). For the adaptation from REFUGE to PrivateDataset2, the most obvious difference is the style of the fundus images. After the domain adaptation, the style of REFUGE is similar to the source domain and the data distribution has also changed greatly (shown in Fig. 7 (b)) and become closer to the source domain. For PrivateDataset1, image-level features change very little, but the distribution (shown in Fig. 7 (c)) also shows that our method can transform the target domain into a source-like domain without much modification to the original images.

b) Quantitative Evaluation: To quantify our generated source-like images and the ability to capture the internal statistics of the source images, we use Accuracy of diagnosis, Sensitivity and Specificity value as the evaluation criteria system. The sensitivity and specificity value are important evaluation index in glaucoma diagnosis tasks based on fundus images. The sensitivity indicates the proportion of patients who are actually diagnosed and correctly diagnosed. The specificity indicates the proportion of patients who are actually disease-free and diagnosed as disease-free.

The experimental result is shown in Table II. For Drishti-GS to PrivateDataset2, the accuracy rate, the sensitivity and specificity value are greatly improved. Using the DAGD method, the accuracy rate increases from 56.25% to 89.58%, and sensitivity increases from 63.64% to 75.0%. The specificity has also been greatly improved from 54.05% to 94.44%. Compared with the other two datasets, the Drishti-GS dataset is significantly different from PrivateDataset2, but our model performs well and successfully transforms the target domain to the source domain. For the REFUGE dataset, the brightness of images is higher than PrivateDataset2, but we still improve the accuracy, sensitivity and specificity value. The accuracy improves from 88.75% to 89.75%, sensitivity increases from 70.0% to 72.5%, and specificity enhances from 90.83% to 91.66%. For the PrivateDataset1, we increase the accuracy and sensitivity from 79.00% to 82.00% and from 78.13% to 81.25% respectively. The specificity is improved from 79.41% to 82.35%. In three datasets, although the data has a homogeneous domain shift caused by different devices and collection rules, our method still achieves good results.

V. CONCLUSION

In this paper, we proposed DAGD, a GAN-based domain adaptation method for glaucoma diagnosis. The framework includes a UDAGAN model, which can learn the latent feature from target domain images and extract style feature from the source domain, and a classification model, which is trained based on ResNet-50. In addition, we adopt a reconstruction loss, which not only reduces model complexity and model parameters but also keeps the optic disc and label unchanged. The experiment results show that our method performs very well in the domain adaptation task of glaucoma diagnosis. Future work will focus on advancing GANs such as pix2pixHD for high-resolution fundus image synthesis and further attempts to generate source-like images without changing many pixel-level features.

REFERENCES

[1] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," in *International conference on machine learning*, 2014, pp. 647–655.

[2] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" *Advances in Neural Information Processing Systems (NIPS)*, pp. 3320–3328, 2014.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[4] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. Vaughan, "A theory of learning from different domains," *Machine Learning*, vol. 79, pp. 151–175, 05 2010.

[5] A. Torralba and A. A. Efros, "Unbiased look at dataset bias," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011.

[6] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," 2014.

[7] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," *arXiv preprint arXiv:1502.02791*, 2015.

[8] B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," *CoRR*, vol. abs/1511.05547, 2015.

[9] B. Sun and K. Saenko, "Deep CORAL: correlation alignment for deep domain adaptation," *CoRR*, vol. abs/1607.01719, 2016.

[10] M. Ghifary, W. B. Kleijn, M. Zhang, D. Balduzzi, and W. Li, "Deep reconstruction-classification networks for unsupervised domain adaptation," in *European Conference on Computer Vision*. Springer, 2016, pp. 597–613.

[11] A. Udayakumar, "Generative adversarial network (gan)," *Artificial Intelligence*, vol. 1, p. 6, 02 2019.

[12] E. Tzeng, J. Hoffman, T. Darrell, and K. Saenko, "Simultaneous deep transfer across domains and tasks," in *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[13] M.-Y. Liu and O. Tuzel, "Coupled generative adversarial networks," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Curran Associates, Inc., 2016, pp. 469–477. [Online]. Available: http://papers.nips.cc/paper/6544-coupled-generative-adversarial-networks.pdf

[14] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," *arXiv preprint arXiv:1409.7495*, 2014.

[15] A. Gretton, A. Smola, J. Huang, M. Schmittfull, K. Borgwardt, and B. Schölkopf, *Covariate shift and local learning by distribution matching*. Cambridge, MA, USA: MIT Press, 2009, pp. 131–160.

[16] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based generative adversarial network," *arXiv preprint arXiv:1609.03126*, 2016.

[17] E. Denton, S. Chintala, A. Szlam, and R. Fergus, "Deep generative image models using a laplacian pyramid of adversarial networks," *CoRR*, p. 1486–1494, 2015.

[18] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *Computer Science*, 2015.

[19] A. Diaz-Pinto, A. Colomer, V. Naranjo, S. Morales, Y. Xu, and A. Frangi, "Retinal image synthesis and semi-supervised learning for glaucoma assessment," *IEEE Transactions on Medical Imaging*, vol. PP, pp. 1–1, 03 2019.

[20] J. Choi, T. Kim, and C. Kim, "Self-ensembling with gan-based data augmentation for domain adaptation in semantic segmentation," in *The IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[21] S. Sankaranarayanan, Y. Balaji, C. D. Castillo, and R. Chellappa, "Generate to adapt: Aligning domains using generative adversarial networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8503–8512, 2018.

[22] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[23] P. Russo, F. M. Carlucci, T. Tommasi, and B. Caputo, "From source to target and back: Symmetric bi-directional adaptive gan," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[24] Y. Zhang, S. Miao, T. Mansi, and R. Liao, "Task driven generative modeling for unsupervised domain adaptation: Application to x-ray image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 599–607.

[25] C. Chen, Q. Dou, H. Chen, and P.-A. Heng, "Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest x-ray segmentation," in *International Workshop on Machine Learning in Medical Imaging*. Springer, 2018, pp. 143–151.

[26] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Krishnan, "Unsupervised pixel-level domain adaptation with generative adversarial networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017, pp. 95–104.

[27] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. A. Efros, and T. Darrell, "Cycada: Cycle-consistent adversarial domain adaptation," *arXiv preprint arXiv:1711.03213*, 2017.

[28] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[29] H. Zhao, H. Li, S. Maurer-Stroh, Y. Guo, Q. Deng, and L. Cheng, "Supervised segmentation of un-annotated retinal fundus images by synthesis," *IEEE Transactions on Medical Imaging*, vol. PP, pp. 1–1, 07 2018.

[30] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[31] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," *arXiv preprint arXiv:1502.02791*, 2015.

[32] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, *Domain-adversarial training of neural networks*. JMLR. org, 2016, vol. 17, no. 1.

[33] C. Chen, Q. Dou, H. Chen, J. Qin, and P.-A. Heng, "Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 865–872, 07 2019.

[34] X. Huang and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," in *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[35] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[36] C. Li and M. Wand, "Precomputed real-time texture synthesis with markovian generative adversarial networks," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 702–716.

[37] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[38] G. G. Chrysos, J. Kossaifi, and S. Zafeiriou, "RoCGAN: Robust conditional GAN," in *International Conference on Learning Representations*, 2019. [Online]. Available: https://openreview.net/forum?id=Byg0DsCqYQ

[39] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 5767–5777. [Online]. Available: http://papers.nips.cc/paper/7159-improved-training-of-wasserstein-gans.pdf

[40] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," *arXiv preprint arXiv:1701.07875*, 2017.

[41] L. van der Maaten and G. Hinton, "Viualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 11 2008.

[42] J. Sivaswamy, S. R. Krishnadas, G. Datt Joshi, M. Jain, and A. U. Syed Tabish, "Drishti-gs: Retinal image dataset for optic nerve head(onh) segmentation," in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, April 2014, pp. 53–56.

[43] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[44] H. Fu, J. Cheng, Y. Xu, C. Zhang, D. Wong, J. Liu, and X. Cao, "Disc-aware ensemble network for glaucoma screening from fundus image," *IEEE Transactions on Medical Imaging*, vol. PP, pp. 1–1, 05 2018.

[45] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 12 2014.