Generative Adversarial Learning for Reducing Manual Annotation in Semantic Segmentation on Large Scale Miscroscopy Images: Automated Vessel Segmentation in Retinal Fundus Image as Test Case

Avisek Lahiri[1]   Kumar Ayush[2]   Prabir Kumar Biswas [3]   Pabitra Mitra[4]

Indian Institute of Technology Kharagpur

West Bengal 721302, India

{[1]avisek, [3]pkb }@ece.iitkgp.ernet.in, [2]kumar.ayush@iitkgp.ac.in, [4]pabitra@cse.iitkgp.ernet.in

Abstract

Convolutional Neural Network(CNN) based semantic segmentation require extensive pixel level manual annotation which is daunting for large microscopic images. The paper is aimed towards mitigating this labeling effort by leveraging the recent concept of generative adversarial network(GAN) wherein a generator maps latent noise space to realistic images while a discriminator differentiates between samples drawn from database and generator. We extend this concept to a multi task learning wherein a discriminator-classifier network differentiates between fake/real examples and also assigns correct class labels. Though our concept is generic, we applied it for the challenging task of vessel segmentation in fundus images. We show that proposed method is more data efficient than a CNN. Specifically, with 150K, 30K and 15K training examples, proposed method achieves mean AUC of 0.962, 0.945 and 0.931 respectively, whereas the simple CNN achieves AUC of 0.960, 0.921 and 0.916 respectively.
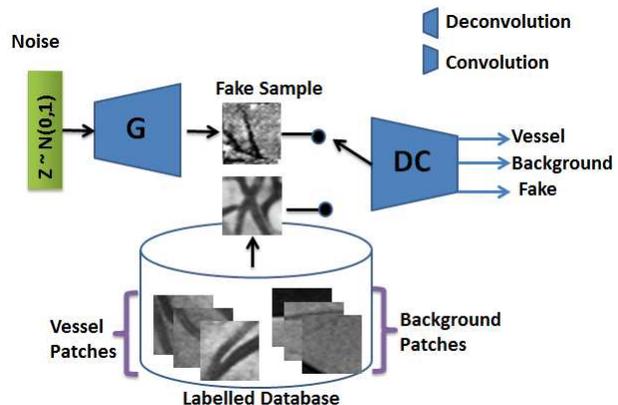
Figure 1. Proposed model for GAN based semantic segmentation on fundus images. The generator network (G) takes in a 300-D standard normal noise vector to create a fake example, G(z) via a series of deconvolution [25] operations. The task of the Discriminator-Classifier network (DC) is to assign correct class label (vessel or background) to real examples coming from stored training database while assigning G(z) to Fake class. In contrast, goal of G is to fool DC in assigning G(z) to any one of the training labels.

1. Introduction

Semantic segmentation refers to pixel level assignment of class labels. The task is particularly more difficult than object detection because a semantic segmenter is penalized even for misclassifying a single pixel. With the advent of deep learning paradigms, semantic segmentation has reached new benchmarks in natural scene understanding [17, 25], large scale medical imaging[31] and microscopic images [4, 30, 38]. But, pixel level manual annotation is a costly and time consuming process. The paper is specifically targeted towards mitigating this extensive labeling bottleneck via a multi task setting using generative adversarial networks (GAN).

Fundus imaging (See Fig. 2) is a paradigm of photographing the internal surface of eye including retina, reti-

nal vasculature, optic disc, macula, and posterior pole or the fundus. A fundus camera is a specialized low power microscope with an attached camera and operates on the principle of indirect ophthalmoscope [40]. Fundus angiograpgy or study of blood vessel network is a key step in detection of diabetic retinopathy [37], localizing of foveal avascular region [8], thinning of arteries [7] and laser surgery [11]. The vascular network also finds application in biometrics [21]. Automated segmentation of blood vessels from fundus image is a challenging task due to intricate branching patterns, noisy background and illumination difference. Thus automated vessel segmentation as a first step toward computer aided analysis of fundus still remains an active research [5, 20] avenue. Though automated vessel segmenta-

tion in fundus images was a well researched topic from last decade [22, 28, 24], the algorithms depended on indigenous hand crafted features. But, in recent years, the paradigm has shifted towards data driven learning by leveraging hierarchical feature representation capability of deep neural networks. Algorithms exploiting deep autoencoders and convolutional neural networks (CNN) has set up new benchmarks for retinal vessel segmentation [5, 13, 20]. There are broadly two paradigms of applying deep learning algorithms for vessel segmentation, viz., a) **Central pixel prediction:** In this method, a training algorithm is fed with a fixed sized rectangular image patch and label of central pixel is assigned as the class label of the patch. Some of the initial deep learning algorithms for this task used deep stacked denoising auto encoder for reconstructing patches extracted from training images. This is a stage wise unsupervised learning framework in which the neural network greedily learns to efficiently encode fundus patches into lower dimensional space. Next step is to refine the entire network end-to-end with class labels. Recent works by [18, 13, 32] closely follows this method. Authors have also used end-to-end supervised training of CNNs on the image patches [19]. In this paper we focus on central pixel prediction. b) **Structure prediction:** Recent works on semantic segmentation for natural images show that it is more beneficial to predict class labels for all pixels of a given patch rather than the central pixel. This is poplar in computer vision community as deep structural learning [15] since it aids in better encapsulation of neighboring structure information for making class predictions. These models are also popular as fully convolutional semantic segmenters (FCN) [17]. Such FCNs have been successfully applied for retinal vessel segmentation to achieve benchmark performances [1].

But training a CNN is data extensive task. This paper aims at reducing manual annotation effort by posing the segmentation problem in a semi supervised learning framework which leverages using the recent concept of generative adversarial network [6]. It is to be noted that the contribution of the paper is not in achieving new state-of-the-art performance in vessel segmentation but rather to showcase the data efficacy of GAN based semantic segmentation paradigm. Though the concept is generic, we apply the proposed method for label free angiography on fundus images. Our main contributions in this work are summarized below:

- To our best knowledge, this is the fist work which leverages GAN for semi supervised learning on large scale fundus imaging modality for automated blood vessel segmentation.

- We achieve comparable performance (sometimes even better) with recent CNN based segmentation techniques while using upto 9X times less training data

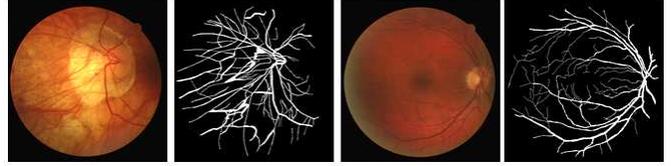- We show that performance of simple CNN based seg-

Figure 2. Exemplary training fundus images with corresponding manual annotation of blood vessels from DRIVE dataset. The task of automated vessel segmentation is particularly challenging due to significant illumination variation, abrupt torosity of vessels and complex vascular network.

menter starts deteriorating faster on smaller datasets compared to GAN-CNN

- For the first time, we generate 32×32 dimensional synthetic fundus patches from latent noise vectors. The synthetic samples closely resemble the original training examples.

- We show that the difference of performances between simple CNN and GAN-CNN is statistically significant when trained on smaller training sets.

2. Generative Adversarial Learning

Generative adversarial network (GAN) [6] presents a two player min-max game between a generator (G) and discriminator (D) network. The idea is to simultaneously train the D and G networks. G is trained to map random vectors $\mathbf{z} \in \mathbb{R}^Z$ to synthetic image vector, $\tilde{\mathbf{x}} = G(\mathbf{z})$. The objective of D network is to distinguish between real examples, $\mathbf{x} \sim p_{data}(x)$, from synthetic examples, $G(z) \sim p_G(z)$ generated by G. $D(\mathbf{x})$ represents the probability that a sample $\mathbf{x}$ belongs to original data distribution. Gradient of output of D with respect to its input is used by G to update its own parameters. Specifically, D and G play a two player min-max game with value function V(G,D):

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

Though, initially GAN was proposed as an approximate sampler from original data distribution, the concept of adversarial learning have been successfully applied over diversified computer vision applications such as image super resolution [14], image impainting [41], image-to-image translation [16] and video frame prediction [23] to name a few.

3. Proposed Method

3.1. Semi Supervised Learning with GAN

The original version of GAN can be implemented with 2-way softmax output from discriminator network to find

a distribution over [REAL, FAKE]. For the task of semi supervised learning on a database of K classes, the concept can be easily extended to incorporate a $K + 1$-way softmax layer at output of discriminator [26, 34] ; where, now the prediction labels will be [Class 1, Class, 2,.., Class K, FAKE]. The revised Discriminator can be termed as a Discriminator-Classifier network (DC net). Let, $p_{DC}(c = k|x)$ define the probability of belonging to class k given an example x by the DC net. k=1,2,...K denotes labels from database (in our case K=2; vessel patch and background patch) and k=K+1 denotes the FAKE class. So, $p_{DC}(c = K + 1|x)$ denotes the probability that x is fake corresponding to 1-D(G(z)) in Eq. 1. The DC net now has to minimize two types of losses, viz. a) classification loss and b)adversarial loss. Classification loss, L_c is given by:

$$L_c = -\mathbb{E}_{(x,y)\sim p_{data}(x,y)}log\ p_{DC}(y|x; y < K + 1) \quad (2)$$

Adversarial loss, L_{avd}, is given by,

$$L_{adv} = -\mathbb{E}_{x \sim G}log\ p_{DC}(y = K + 1|x) \quad (3)$$

The DC net is optimized to minimize Eqs. 2 and 3. The generator is updated in such a way so that the DC net places minimum probability over class k=K+1 and thereby fooling DC net to believe that the fake example belongs to one of the legitimate K classes of the database. So, for training the generator, we need to maximize, L_G,

$$L_G = -\mathbb{E}_{x \sim G}log\ p_{DC}(y = K + 1|x) \quad (4)$$

At test time, a real test examples, x_t is assigned a label, $y^*(x)$, according to,

$$y^*(x) = \underset{y}{\mathrm{argmax}}\ p_{DC}(c = y|x) \quad (5)$$

The basic diagram of our method is shown in Fig. 1.

DC net can be seen as a multi task learning network [39, 2] where the two tasks of the network is to assign correct class label to a real training sample, x $\sim p_{data}(x)$, while to assign high probability to FAKE class when fed with synthetic example, $G(z)$. It has been shown in numerous computer vision applications, that multi task learning aids in improving performance over individual tasks provided that the tasks are related [3]. In [34], the authors achieve state-of-the-art semi supervised performance on MNIST and CIFAR datasets using GAN framework.

Invigorated by the success of GAN in semi supervised learning setting, we incorporate a similar learning strategy for semantic segmentation for large scale fundus images. Pixel level annotation is much tedious than image level tagging; thereby bolstering the importance of our contribution towards data efficient semantic segmentation modeling.

3.2. Network Architecture

The network architecture is visualized in Fig. 1. We closely adhere to the GAN network specifications of DCGAN [29] [1]. Noise vector, z is randomly sampled from 300-D zero mean unit variance normal distribution. z is then projected to a 32,768-D space and transformed back to a 8×8 cuboid with 512 channels. Three stages of deconvolution are used for upsampling from 8×8 to 64×64 dimension. Though batch normalization [9] was initially suggested in DCGAN, recently, it has been observed that batch normalization introduces irrelevant stochasticity for small batch sizes. Following [10], we apply instance normalization after every deconvolution followed by Rectified linear unit (ReLU) as non linear activation.

The DC network receives a 64×64 image and passes it through a series of leaky ReLU activated convolutional layers with instance normalization. At last layer, it produces a 3-way softmax output corresponding to background, vessel and fake respectively. Let CILR(f,r,s) denote a convolutional layer with leaky Relu activation after instance normalization. The layer operates with a receptive field of r×r with stride of s along height and width and yields f filters. The architecture of the DC network is as follows: CILR(96,3,1)-CILR(96,3,1)-CILR(96,3,2)-dropout(0.5)-CILR(192,3,1)-CILR(192,3,1)-CILR(192,3,2)-dropout(0.5)-CILR(192,3,1)-CILR(192,1,1)-global_avg-softmax(3); where dropout(k) is the standard dropout layer for regularization [35] with keep node probability of k and global_avg is global average along height and width dimension followed by 3-way softmax layer.

4. Experiments

In this section we provide the details of training dataset, training procedures and provide comparative analysis to show the effectiveness of using proposed GAN-CNN for semantic segmentation over simple CNN.

4.1. Fundus Dataset

The experiments have been performed on the fundus images of DRIVE dataset [36]. The dataset contains 20 images for training and 20 for testing. Blood vessel in each image is manually marked by an observer trained by experienced ophthalmologist. This marking is taken as gold standard. For testing images, there was an additional human marker to compare efficacy of automated vessel segmentation algorithms with human perfection.

4.2. Preprocessng

It has been shown in [13], that then green channel in color fundus imaging is most discriminative in segmenting

---

[1]Our code is adapted from https://github.com/carpedm20/DCGAN-tensorflow

blood vessels. Following this, we considered only the green channel and performed histogram equalization to mitigate effects of non uniform illumination. Next, 64×64 dimensional patches were extracted and label of central pixel is assigned as the class label of the entire patch. Following [13], vessel patches are extracted from skeletonized ground truth maps while background patches are extracted from dilated maps. This helps in better prediction of pixels at the border of vessels.

Table 1. Comparison of mean AUC of proposed GAN-CNN and simple CNN based segmentation on test set of DRIVE dataset. Comparison is done after training on different sized dataset. p-value (Welch's t-test [33]) indicates that there is significant difference between the mean AUCs of the proposed method and simple CNN, specially when trained on smaller training sets.

Dataset	GAN-CNN(Proposed)	CNN	p-value
150K	0.962	0.960	0.1
30K	0.945	0.921	10^{-3}
15K	0.931	0.916	10^{-5}

4.3. Training Details

For understanding the difference of performance between proposed GAN-CNN and simple CNN based method, we train on three different sized dataset, viz., a)150K b) 30K c) 15K consisting on 150×10^3, 30×10^3 and 15×10^3 training patches respectively. On each dataset, we train the simple CNN and GAN-CNN from scratch.

ADAM optimizer[12] is used for updating both the G and DC net. Initial learning rate for both G and DC is kept at 10^{-4} with a decay factor of 0.8 after every 20 epochs. Slope of leaky ReLU was maintained at 0.1. Trying to optimize the generator network with Eq. 4 is practically not advisable [6] because in the early phase of training, magnitudes of gradients propagated to generator are small. Thus, we instead minimize,

$$L_G = -\mathbb{E}_{x \sim G} log\{1 - p_{DC}(y = K + 1|x)\} \quad (6)$$

4.4. Results

In retinal vessel segmentation literature, area under the Receiver Operation Curve, i.e., AUC is taken as a standard metric of comparison [5, 18, 19]. A larger AUC signifies a better segmenter. First, we compare the data efficiency of our proposed GAN-CNN and simple CNN based segmenter. In Table 1, we report the mean AUC of the two methods trained on 150K, 30K and 15K datasets and tested on the 20 test images. With 150K training, performances of both GAN-CNN and simple CNN are comparable with mean AUC values of 0.962 and 0.960 respectively. But, the superiority of GAN-CNN is manifested with reduction of training data. At 30K training, mean AUC of GAN-CNN is 0.9405 while that of simple CNN is 0.921; at 15K training,

Table 2. Comparison of mean AUC of some of the contemporary deep learning based retinal vessel segmentation algorithm. Our proposed methods manifest comparable performance even when trained with much smaller training datasets. It is to be noted that in this preliminary work we have not performed extensive grid search for best learning rate scheduling or architecture fine tuning. So, there is future scope to enhance the data efficiency of our model even further.

Method	Dataset Size	AUC
Maji et al. [18]	60K	0.928
Lahiri et al. [13]	120K	0.950
Fu et al.[5]	330K	0.947
Maji et al. [19]	60K	0.919
Liskowski et al. [1]	3857K	0.963
GAN-CNN (**Proposed**)	30K	0.945
GAN-CNN (**Proposed**)	15K	0.931

the corresponding values are 0.931 and 0.916 respectively. We also visualize the ROC curves of the two methods on the entire test dataset in Fig. 3. It can be seen that the ROC curve of proposed GAN-CNN is always higher than simple CNN when trained on smaller scale dataset. Table 1 and Fig. 3 thus strongly advocates use of GAN based segmentation models specially while training on limited amount of labeled data.

Table 1 shows that the mean AUC of GAN-CNN is greater than that of simple CNN. To test the significance of this observation we perform a Welch's t-test which is a special adaptation of Student's t-test for unequal variances [33]. The null hypothesis in this case is that the mean AUCs of both paradigms of segmenters are same. The p-value for Welch's t-test comes out to be 0.1, 10^{-5} and 10^{-7} on 150K, 30K and 15K training set respectively. This shows that the difference of mean AUCs between GAN-CNN and simple CNN is significant specially when trained on limited dataset. For the 150K dataset training, differences in performance of GAN-CNN and CNN are statistically insignificant.

We also compare the performance of our model trained on limited data with some of the contemporary deep learning based techniques for automated retinal vessel segmentation. The competing methods are a) Maji et al. [18] : The authors first train a denoised stacked auto encoder for vessel reconstruction followed by supervised refinement of the autoencoder network. The learnt features are then fed to an ensemble of random forest for final classification; b) Lahiri et al. [13]: Here, a 2-stage ensemble (based on bootstrap sampling and architectural variation) of stacked denoised autoencoder is used for classifying a central pixel to either vessel or background class; c) Liskowski et al. [1]: An end-to-end CNN based network is proposed with different variants of image preprocessing and structured predic-
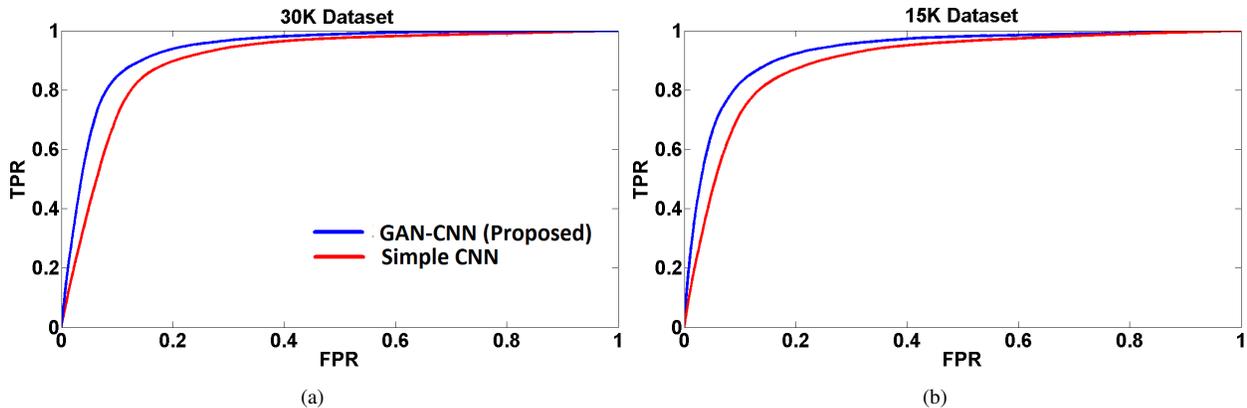
Figure 3. ROC curves of proposed GAN-CNN and simple CNN on the combined 20 test images of DRIVE retina dataset. 30K and 15K dataset contains total 30,000 and 15,000 training patches respectively. Curves of GAN-CNN always tends to be higher on the ROC plots compared to simple CNN based segmenter. The visualization bolsters our claim that training a GAN based CNN for semantic segmentation is data efficient.

tion; d) Fu et al. [5]: The authors train a deep CNN followed by finer refinement using conditional random fields e) Maji et al. [19]: An ensemble of deep CNNs is used for pixelwise prediction. We tabulate the comparative performances in Table 2. It is to be noted that even with much smaller dataset size, our proposed method performs comparable (sometimes even better) than the competing techniques trained with 2X-10X times more training data. Also, in this preliminary work, we have not done extensive grid search for hyper parameter tuning or architectural refinement. In a follow up work we wish to perform a detailed refinement with an envision of further enhancement of data efficiency of our model.

In Fig. 4, we visualize an exemplary set of examples generated by the generator network of our GAN based model and real training patches from the database. It can be appreciated that the parameterization of the trained generator enables it to map a 300-D latent space to viable 64×64 dimensional fundus patches. Specially, the anatomy of vessel representation learnt by the generator is worth considering.

5. Conclusion

In this paper we proposed a semi supervised paradigm of semantic segmentation using generative adversarial networks. The training of the proposed method was shown to be more data efficient compared to normal CNN training. Our preliminary work thus advocates future researchers working on large scale microscopic images to leverage GAN based CNN model for semantic segmentation to reduce manual labeling effort. There are some immediate extensions possible from the presented work. One possibility is to make use of large amount of unlabeled data by forcing the DC-net to place low likelihood for fake class to these ex-

amples [26]. Another possibility to use use class conditional generator [27] network to force it to generate class specific fake examples and forcing the DC-net to classify these fake examples. Both of these methods are further steps towards improving the performance of the combined DC-net.

References

[1] Segmenting retinal blood vessels with deep neural networks. 2, 4

[2] A. Argyriou, T. Evgeniou, and M. Pontil. Multi-task feature learning. *Advances in neural information processing systems*, 19:41, 2007. 3

[3] R. Caruana. Multitask learning. In *Learning to learn*, pages 95–133. Springer, 1998. 3

[4] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pages 2843–2851, 2012. 1

[5] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, pages 698–701. IEEE, 2016. 1, 2, 4, 5

[6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014. 2, 4

[7] E. Grisan and A. Ruggeri. A divide et impera strategy for automatic classification of retinal vessels into arteries and veins. In *Engineering in medicine and biology society, 2003. Proceedings of the 25th annual international conference of the IEEE*, volume 1, pages 890–893. IEEE, 2003. 1

[8] A. Haddouche, M. Adel, M. Rasigni, J. Conrath, and S. Bourennane. Detection of the foveal avascular zone on retinal angiograms using markov random fields. *Digital Signal Processing*, 20(1):149–154, 2010. 1
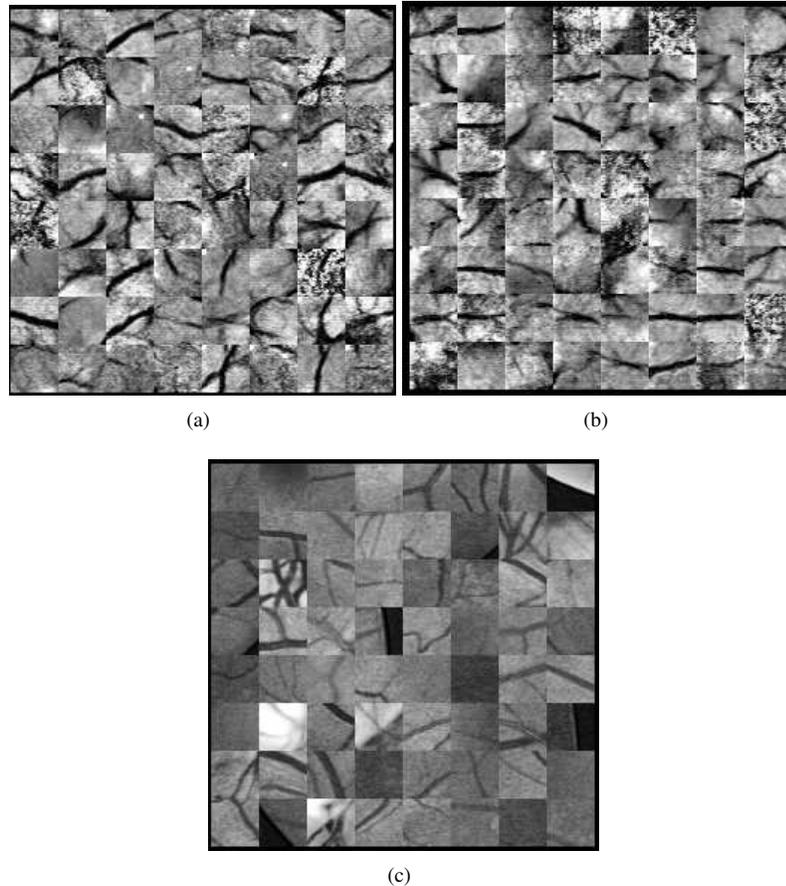
Figure 4. Exemplary real and generated image patches. (a): Samples generated during training on 30K dataset; (b): Samples generated during training on 15K dataset; (c) Real samples for training set.

[9] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015. 3

[10] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer, 2016. 3

[11] J. J. Kanski and B. Bowling. *Clinical ophthalmology: a systematic approach.* Elsevier Health Sciences, 2011. 1

[12] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 4

[13] A. Lahiri, A. G. Roy, D. Sheet, and P. K. Biswas. Deep neural ensemble for retinal vessel segmentation in fundus images towards achieving label-free angiography. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 1340–1343. IEEE, 2016. 2, 3, 4

[14] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. *arXiv preprint arXiv:1609.04802*, 2016. 2

[15] G. Lin, C. Shen, A. van den Hengel, and I. Reid. Efficient piecewise training of deep structured models for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3194–3203, 2016. 2

[16] M.-Y. Liu, T. Breuel, and J. Kautz. Unsupervised image-to-image translation networks. *arXiv preprint arXiv:1703.00848*, 2017. 2

[17] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015. 1, 2

[18] D. Maji, A. Santara, S. Ghosh, D. Sheet, and P. Mitra. Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 3029–3032. IEEE, 2015. 2, 4

[19] D. Maji, A. Santara, P. Mitra, and D. Sheet. Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images. *arXiv preprint arXiv:1603.04833*, 2016. 2, 4, 5

[20] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool. Deep retinal image understanding. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 140–148. Springer, 2016. 1, 2

[21] C. Mariño, M. G. Penedo, M. Penas, M. J. Carreira, and F. Gonzalez. Personal authentication using digital retinal images. *Pattern Analysis and Applications*, 9(1):21, 2006. 1

[22] M. E. Martinez-Perez, A. D. Hughes, S. A. Thom, A. A. Bharath, and K. H. Parker. Segmentation of blood vessels from red-free and fluorescein retinal images. *Medical image analysis*, 11(1):47–61, 2007. 2

[23] M. Mathieu, C. Couprie, and Y. LeCun. Deep multi-scale video prediction beyond mean square error. *arXiv preprint arXiv:1511.05440*, 2015. 2

[24] A. M. Mendonca and A. Campilho. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE transactions on medical imaging*, 25(9):1200–1213, 2006. 2

[25] H. Noh, S. Hong, and B. Han. Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1520–1528, 2015. 1

[26] A. Odena. Semi-supervised learning with generative adversarial networks. *arXiv preprint arXiv:1606.01583*, 2016. 3, 5

[27] A. Odena, C. Olah, and J. Shlens. Conditional image synthesis with auxiliary classifier gans. *arXiv preprint arXiv:1610.09585*, 2016. 5

[28] R. Perfetti, E. Ricci, D. Casali, and G. Costantini. Cellular neural networks with virtual template expansion for retinal vessel segmentation. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 54(2):141–145, 2007. 2

[29] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015. 3

[30] S. Ragothaman, S. Narasimhan, M. G. Basavaraj, and R. Dewar. Unsupervised segmentation of cervical cell images using gaussian mixture model. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 70–75, 2016. 1

[31] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015. 1

[32] A. G. Roy and D. Sheet. Dasa: Domain adaptation in stacked autoencoders using systematic dropout. In *Pattern Recognition (ACPR), 2015 3rd IAPR Asian Conference on*, pages 735–739. IEEE, 2015. 2

[33] G. D. Ruxton. The unequal variance t-test is an underused alternative to student's t-test and the mann–whitney u test. *Behavioral Ecology*, 17(4):688–690, 2006. 4

[34] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. Improved techniques for training gans. In *Advances in Neural Information Processing Systems*, pages 2226–2234, 2016. 3

[35] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1):1929–1958, 2014. 3

[36] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging*, 23(4):501–509, 2004. 3

[37] T. Teng, M. Lefley, and D. Claremont. Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy. *Medical and Biological Engineering and Computing*, 40(1):2–13, 2002. 1

[38] K. Xu, H. Su, J. Zhu, J.-S. Guan, and B. Zhang. Neuron segmentation based on cnn with semi-supervised regularization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 20–28, 2016. 1

[39] Y. Xue, X. Liao, L. Carin, and B. Krishnapuram. Multitask learning for classification with dirichlet process priors. *Journal of Machine Learning Research*, 8(Jan):35–63, 2007. 3

[40] K. Yang-Williams. Ophthalmic photography: Retinal photography, angiography, and electronic imaging, 2002. 1

[41] R. Yeh, C. Chen, T. Y. Lim, M. Hasegawa-Johnson, and M. N. Do. Semantic image inpainting with perceptual and contextual losses. *arXiv preprint arXiv:1607.07539*, 2016. 2