UWAT-GAN: Fundus Fluorescein Angiography Synthesis via Ultra-wide-angle Transformation Multi-scale GAN

Zhaojie Fang[1] Zhanghao Chen[1] Pengxue Wei[3] Wangting Li[3] Shaochong Zhang[3] Ahmed Elazab[4] Gangyong Jia[2] Ruiquan Ge[2*] Changmiao Wang[5*]

1 HDU-ITMO Joint Institute, Hangzhou Dianzi University, Hangzhou 310018, China
2 School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China
3 Shenzhen Eye Hospital, Jinan University, Shenzhen 518040, China
4 School of Biomedical Engineering, Shenzhen University, Shenzhen 518060, China
5 Medical Big Data Lab, Shenzhen Research Institute of Big Data, Shenzhen 518172, China

*Corresponding authors: gespring@hdu.edu.cn; cmwangalbert@gmail.com

MICCAI 2023 Vancouver CANADA

杭州電子科技大学 HANGZHOU DIANZI UNIVERSITY
深圳市大数据研究院 Shenzhen Research Institute of Big Data
SZEH 深圳市眼科医院 SHENZHEN EYE HOSPITAL
深圳大学 SHENZHEN UNIVERSITY

Abstract

Fundus photography is an essential examination for clinical and differential diagnosis of fundus diseases. Recently, Ultra-Wide-angle Fundus (UWF) techniques, UWF Fluorescein Angiography (UWF-FA) and UWF Scanning Laser Ophthalmoscopy (UWF-SLO) have been gradually put into use. However, Fluorescein Angiography (FA) and UWF-FA require injecting sodium fluorescein which may have detrimental influences. To avoid negative impacts, cross-modality medical image generation algorithms have been proposed. Nevertheless, current methods in fundus imaging could not produce high-resolution images and are unable to capture tiny vascular lesion areas. This paper proposes **a novel conditional generative adversarial network (UWAT-GAN) to synthesize UWF-FA from UWF-SLO.** Using multi-scale generators and a fusion module patch to better extract global and local information, our model can generate high-resolution images. Moreover, an attention transmit module is proposed to help the decoder learn effectively. Besides, a supervised approach is employed to train the network using multiple new weighted losses on different scales of data. **Experiments on an in-house UWF image dataset demonstrate the superiority of the UWAT-GAN over the state-of-the-art methods.** Our source code is available at: **https://github.com/Tinysqua/UWAT-GAN**.

Contributions

⇒ **The first study to synthesize UWF-FA from UWF-SLO**, overcoming the limitations of UWF-FA imaging.

⇒ A novel method UWAT-GAN utilizing **multi-scale generators and multiple new weighted losses** on different data scales synthesizes **high-resolution** images with the ability to **capture tiny vascular lesion areas**.

⇒ Enhance the **clarity of vascular regions** and **tackle the misalignment problems** by adopting an effective pre-processing method for image sharpening and registration.

⇒ The proposed UWAT-GAN **outperforms** the state-of-the-art models through extensive experiments, comparisons, and ablation studies.

Method and Comparisons

We proposed a **supervised conditional GAN** for synthesizing UWF-FA from UWF-SLO images. To achieve this goal, **a fine-coarse level generator that sought to extract local and global information, a fusion module that worked on the outputs of both level generators, and an attention transmit module** were proposed to enhance the U-Net-like architecture, which could be seen in Figure1. Figure 2 showed the visual comparison of our method with current methods containing Cycle-gan and Pi2PixHD. Arrows in Figure 2 pointed to the important areas. It could be seen that Pix2PixHD generated better on macula but worse on optic disc, while Cycle-gan generated poorly on macula. Our approach produced good results both on macula and optic disc. Additionally, we made a comparison with current conditional GAN on **Fréchet Inception Distance (FID), Kernel Inception Distance (KID), Inception Score (IS) and Learned Perceptual Image Patch Similarity (LPIPS), shown in Table 1.** We also compared our method that with and without attention transmit module, and pix2pixHD. The ablation experiment results were visually displayed in Figure 3.

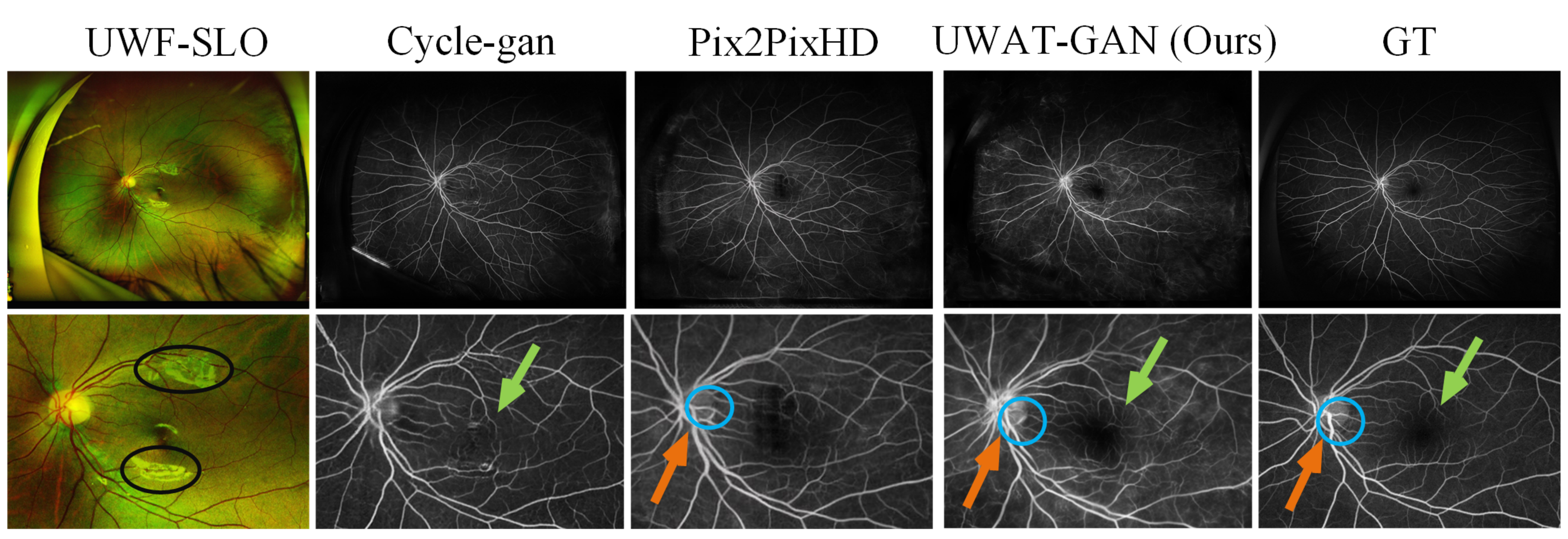

Figure 2. The visual comparison of our method with current methods containing Cycle-gan and Pi2PixHD.

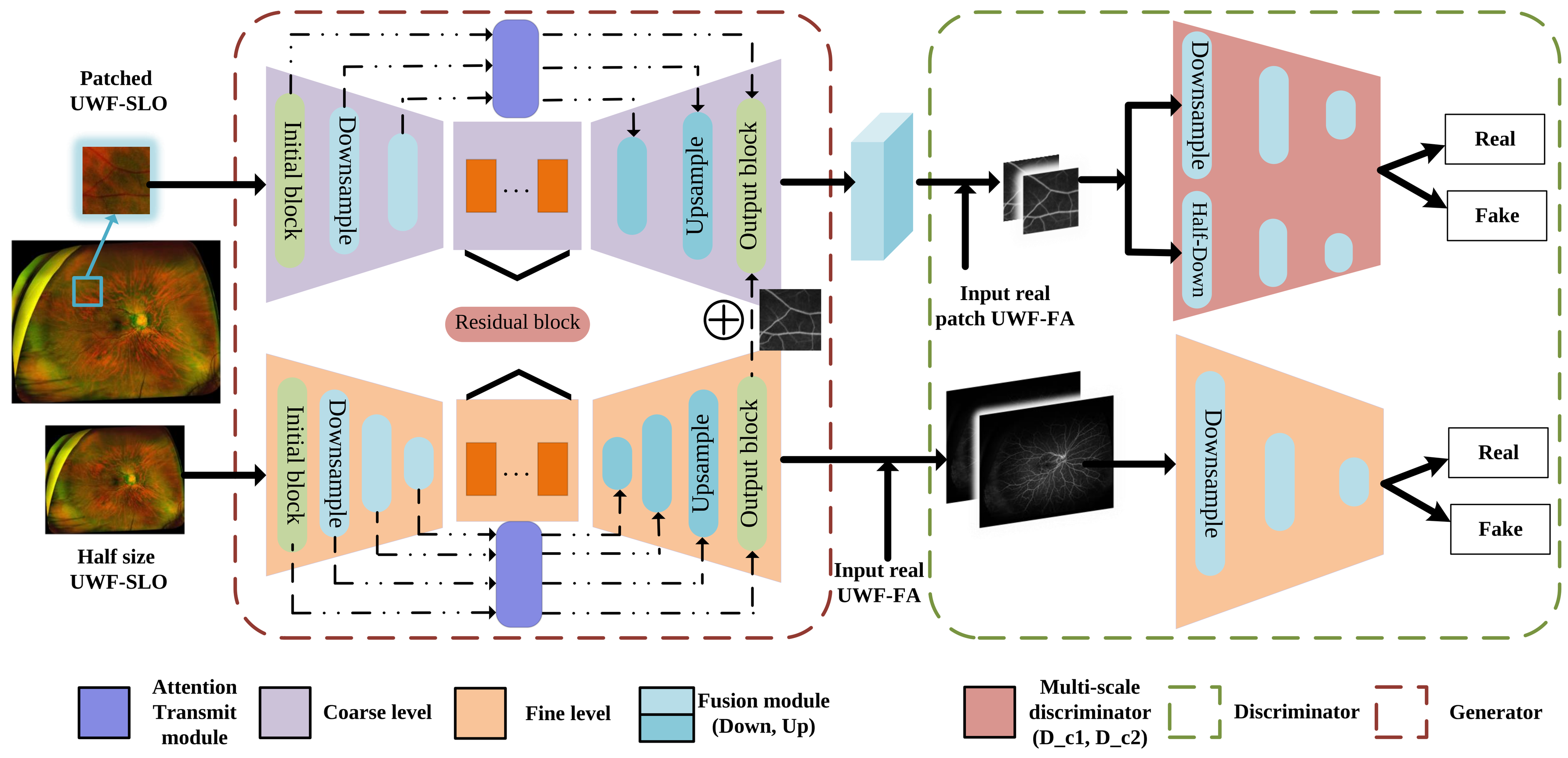

Figure 1. The overall architecture of our method. It contains three modules that we proposed: Fine and Coarse level generator, the multi-scale discriminator and the attention module.

*Table 1. Comparison with the state-of-the-art methods using four evaluation metrics. The * means that the official code has not provide the way to measure it. M_{NA} means without attention transmission module.*

Methods	FID(↓)	KID(↓)	IS(↑)	LPIPS(↓)
Pix2Pix[1]	135.4038	0.1094	1.2772	0.4575
Pix2PixHD[2]	76.76	0.0491	1.0602	0.4451
StarGAN-v2[3]	74.38	0.0433	*	0.4577
UWAT-GAN M_{NA}	67.96	0.0308	1.2757	0.4086
UWAT-GAN(Ours)	**55.59**	**0.0260**	**1.323**	**0.3826**

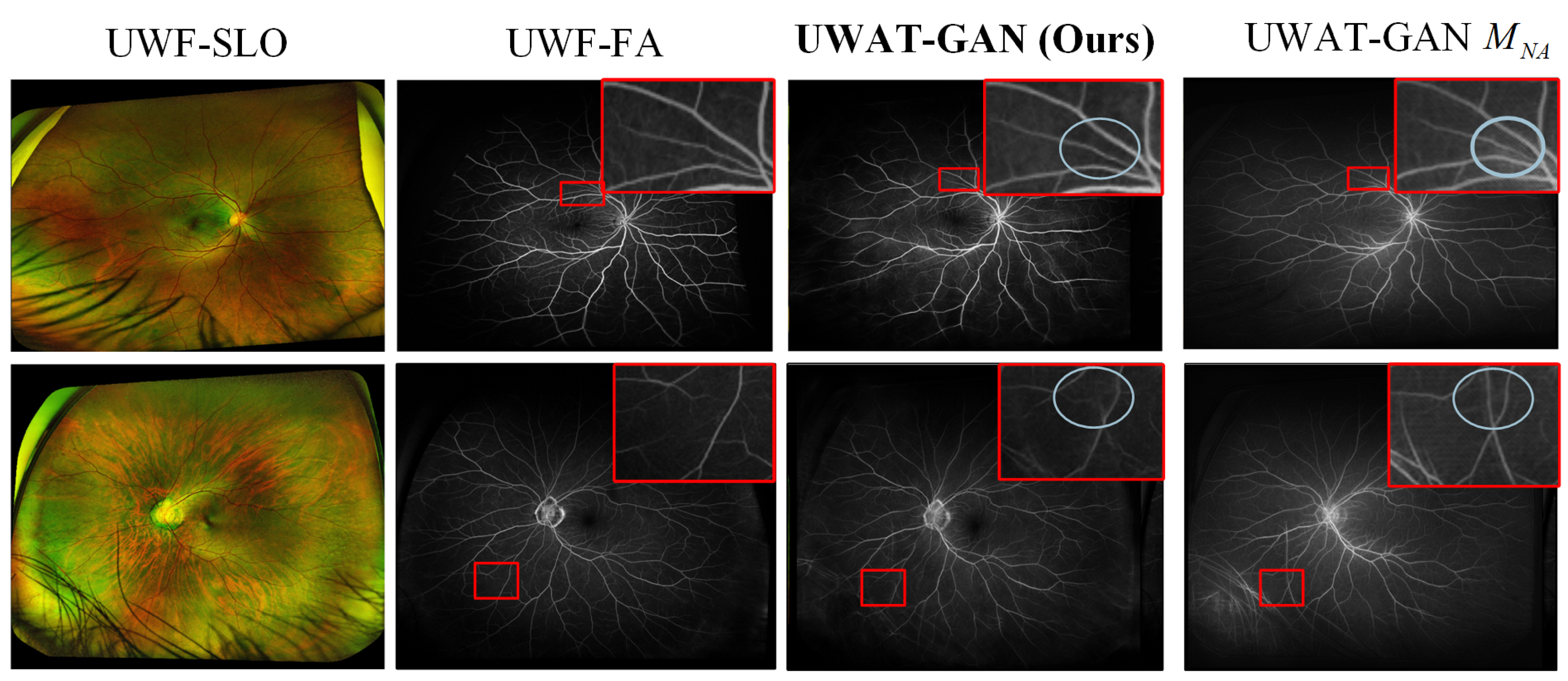

Figure 3. An ablation comparison of our method UWAT-GAN on two pairs of UWF images. From left to right: source UWF-SLO, UWF-FA, the proposed framework with and without the attention transmit module, respectively. We provide the zoom-in view of each UWF-FA image on the top right.

Conclusion

Our work proposed UWAT-GAN to synthesize UWF-FA from UWF-SLO, which could generate **high-resolution images and enhance the ability to capture small vascular lesions**. Comparison and ablation study on an in-house dataset demonstrated the **superiority and effectiveness** of our proposed method.
In the future, we aim to expand the size of our dataset and explore the use of **object detection model**, especially for small targets, to push our model pay more attention to some lesions. After further validation, we aim to adopt this method as **an auxiliary tool** to detect and diagnose fundus diseases.

References

[1] Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1125–1134 (2017)

[2] Wang, T.C., Liu, M.Y., Zhu, J.Y., Tao, A., Kautz, J., Catanzaro, B.: Highresolution image synthesis and semantic manipulation with conditional gans. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 8798–8807 (2018)

[3] Choi, Y., Uh, Y., Yoo, J., Ha, J.W.: Stargan v2: Diverse image synthesis for multiple domains. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8188–8197 (2020)