EDDY WIJANTO-FILE 3

by Eddy Wijanto-file 3 Eddy Wijanto-file 3

Submission date: 14-Feb-2023 02:07PM (UTC+0700)
Submission ID: 2013876146
File name: ical_Coherence_Tomography_System_for_Artifact_Elimination_1.pdf (1.3M)
Word count: 7298
Character count: 38677

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 17, 2021, accepted July 9, 2021, date of publication July 21, 2021, date of current version July 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3098865

Applying a Pix2Pix Generative Adversarial Network to a Fourier-Domain Optical Coherence Tomography System for Artifact Elimination

CHUN-MING HUANG[1], EDDY WIJANTO[2], AND HSU-CHIH CHENG[2]

[1]Department of Electronic Engineering, National Formosa University, Huwei, Yunlin 63201, Taiwan
[2]Department of Electro-Optical Engineering, National Formosa University, Huwei, Yunlin 63201, Taiwan

Corresponding author: Hsu-Chih Cheng (chenghc@nfu.edu.tw)

This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST 108-2221-E-150-041, Grant MOST 107-2218-E-150-008-MY2, and Grant MOST 110-2221-E-150-002.

ABSTRACT The presence of artifacts, including conjugate, DC, and auto-correlation artifacts, is a critical limitation of Fourier-domain optical coherence tomography (FD-OCT). Many methods have been proposed to resolve this problem to obtain high-quality images. Furthermore, the development of deep learning has resulted in many prospective advancements in the medical field; image-to-image translation by using generative adversarial networks (GANs) is one such advancement. In this study, we propose applying the Pix2Pix GAN to eliminate artifacts from FD-OCT images. The first experiment results showed that the proposed framework could translate conventional FD-OCT depth profiles into artifact-free FD-OCT depth profiles. In addition, the FD-OCT depth profile and optical distance of translated images matched those of ground truth images. Second experiment verified that the proposed GAN-based FD-OCT can be applied to generate artifact-free FD-OCT image with different parameters of sample refractive index, the front surface of the sample toward the zero-delay position, and the physical thickness of the sample. Third experiment proved that the proposed model could translated the conventional FD-OCT depth profiles with additional Gaussian noises source image into artifacts-free FD-OCT and successfully relieved the noise.

INDEX TERMS Artifacts, FD-OCT, image-to-image translation, Pix2Pix GAN.

I. INTRODUCTION

Optical coherence tomography (OCT) is an optical imaging modality used to obtain high-resolution cross-sectional tomographic images of the internal microstructures of materials and biological systems. OCT is a noninvasive imaging modality that produces images by using backscattered or back-reflected light. Compared with conventional ultrasound, it can provide higher-resolution images at higher magnitudes and has been widely used for diagnosing ocular diseases. In general, OCT can be classified into two categories: time-domain OCT (TD-OCT) and Fourier-domain OCT (FD-OCT). In FD-OCT, as the signal-to-noise ratio increases in proportional to the number of detection elements, the imaging speed and sensitivity considerably increase significantly. Nevertheless, in a traditional FD-OCT system, since the acquired interferometric signal only represents the real component of a complex wavefo the complex conjugate mirror image is symmetrical to the zero-delay depth. Consequently, because the FD-OCT system is more sensitive around the zero-delay line, imaging is performed by positioning the zero-delay line at the region of interest in a sample to obtain double-depth range images. Sanjay *et al.* [1] used spectral-domain OCT (SD-OCT) to examine patients diagnosed as having glaucoma and reported that 15.2%–36% of scans showed artifacts that may cause difficulty to physicians in the analysis of images. A study identified various types of artifacts that can lead to an incorrect diagnosis [2]. Because artifacts can obscure imaging results and prevent the detection of critical features in a sample structure, full-range FD-OCT is commonly implemented using phase shifting to reconstruct the sample structure and overcome this complex artifact problem. Full-range FD-OCT images are obtained by recording several interferograms with different phase relations. Jiewen *et al.* [3] proposed a five-frame variable phase-shifting (FVP) method to reduce

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaochun Cheng.

the effect of polychromatic errors. Compared with the traditional five-frame invariant phase-shifting method, the FVP method could significantly improve the quality of OCT images with a factor of 1.7 for the suppression of complex conjugate artifacts (CCAs). In our previous studies, we proposed another method for artifact suppression that utilized orthogonal polarized light for phase shifting to improve the speed of image scanning and remove unwanted components [4], [5]. The simulation results of our previous study showed that two FD-OCT interferograms could be simultaneously obtained along with orthogonal polarization components [4]. The benefits of this method were also reported in our further study based on the experimental observation [5]. Furthermore, another study reported that applying an orthogonal interferometer to a nondestructive dimensional metrological system resulted in high speed, high precision, and an ultra-long range [6]; moreover, suppression ratios of 80 dB for direct current (DC) and 60 dB for mirror images could be achieved using this method. Using an ultra-broadband light source to achieve the sinusoidal vibration of a mirror, Qiukun *et al.* [7] obtained a series of spectral interferograms with different phase delays that resulted in the elimination of CCAs. This postprocessing method was also beneficial for increasing the quality of OCT images. Gangjun *et al.* [8] evaluated two methods to align interferograms affected by trigger jitter to reduce residual fixed-pattern noise. The first method involved using a wavenumber shift (k-shift) in the interferograms of interest and searching for the k-shift that minimized the fixed-pattern artifact, whereas the second method involved using relative k-shift and the phase information at the residual fixed-pattern noise location. To remove or reduce artifacts, another study proposed a postprocessing method [9] that involved using a correction factor extracted from a prereconstructed tomogram. Distinguishable morphological features of the sample surface could be detected using this method.

FD-OCT imaging results are highly sensitive to the effects of system instability and environmental noise. This limitation can be resolved using the one-shot phase-shifting method based on a reference wavefront tilting technique. A previous study [10] proposed an achromatic phase-shifting method in which a linear polarizer and a quarter-wave plate were used to generate circularly polarized light in the reference arm. This method could produce fringe-free OCT images in a single shot. However, this method required phase calibration and was hampered by wavelength dispersion effects. Another study proposed an FD-OCT design with two phase-shifted interference fringes that were simultaneously obtained from two orthogonally polarized lights and processed using the image reconstruction algorithm [11]. This proposed FD-OCT design [11] could produce one-shot images by using a fixed apochromatic quarter-wave retarder rather than a rotating polarizer and a mechanical scan. Furthermore, in our previous study, we used one-shot full-range FD-OCT that had a noise elimination feature to measure the thickness and refractive index of samples [12]. The proposed method in [12] provided full-range and one-shot measurements with twice the maximal depth position.

Deep learning (DL) has revolutionized artificial intelligence (AI), and it has solved many complex problems related to AI. DL models are composed of multiple layers, where each layer is connected to its lower and upper layers through different weights. The capability of DL models to learn hierarchical features from various types of data makes them useful for solving many problems, including those encountered in medical imaging. Sripad *et al.* [13] proposed a DL framework to denoise a single-frame OCT B-scan of an optic nerve head (ONH) that provided the advantages of decreased scanning times and minimal patient discomfort. The quantitative measurements of the DL method showed an improvement in the mean signal-to-noise ratio, mean contrast-to-noise ratio, and mean structural similarity index. In addition, the DL method was used to extract capillary-level angiograms from a single OCT volume [14] and detect retinal nerve fiber layer segmentation errors on SD-OCT [15]. Further DL role in medical imaging was showed in [16] for automated age-related macular degeneration (AMD) detection by utilizing support vector machine (SVM), AlexNet, GoogLeNet, and Inception-ResNet for AMD detection while a block-matching and 3-Dimension filter (BM3DF), a hybrid median filter (HMF), and an adaptive wiener filter (AWF) were used to denoise the OCT images. The results reported in [16] indicated that by considering the trade-offs between the computation time and detection accuracy, AlexNet achieves a high detection accuracy with low computation time.

In 2014, a generative adversarial network (GAN) was developed by Goodfellow as a deep-learning-based generative model [17]. The GAN consists of two parts: the generator that learns to generate plausible data, and the discriminator that learns to distinguish the generator's fake data from real data. In the training process, the generator produces fake data, and the discriminator attempts to discriminate between real and fake data. The GAN has shown remarkable results in various tasks such as image generation, image translation, super-resolution imaging, and face image synthesis. Haris *et al.* [18] proposed a single process that could be used to remove both noise and retinal shadows from unseen single-frame B-scans within a short time by using the GAN. Another study used GAN to remove shadows from OCT images, thereby correcting blood vessel shadows in the OCT images of the ONH [19]. In the absence of clean images for training, Guo *et al.* in [20] proposed nonlocal-GAN method for 3D OCT image denoising. The experiment results proved the superiority of the proposed method compare to other denoising approaches while retaining more useful details and clearer layer structures. Further work proposed by Huang *et al.* in [21] with disentangled representation GAN (DRGAN) by utilizing noise and content disentanglement of an OCT image. The noisy image was disentangled into content and noise spaces by corresponding encoders, and then the generator will predict the denoised OCT image through

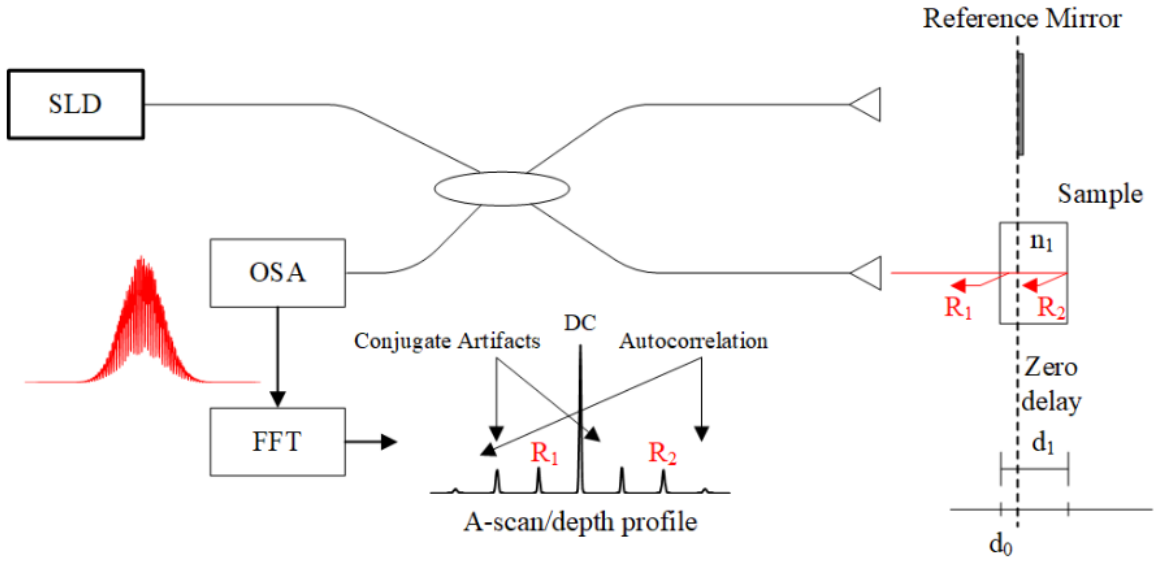

FIGURE 1. Conventional full-range FD-OCT architecture.

these extracted content features. The qualitative and quantitative results from the experiments presented that the proposed method is superior to the other conventional methods. GAN also can be utilized for style transfer, namely noise adaptation GAN (NAGAN) proposed by Zhang *et al.* in [22]. In the proposed scheme, one generator and two discriminators were used for the noise style transfer. The experiment results verified that the noise styles were transferred while maintaining the contents. The application of NAGAN for OCT showed that the noise style transfer is able to improve the segmentation and classification task, both in OCT and ultrasound images, respectively. Furthermore, in a previous study [23], a GAN was implemented with conditional adversarial networks, known as cGAN, to obtain a general-purpose solution. This model can be used to solve various image translation problems.

In this study, the Pix2Pix GAN, as a component of a cGAN framework [23], was applied to the one-shot and full-range FD-OCT system to eliminate artifacts, including conjugate, DC, and auto-correlation artifacts, by using the image-to-image translation mechanism. The Pix2Pix GAN was trained to learn how to translate a conventional FD-OCT depth profile generated from inverse fast Fourier transform (IFFT) signal processing into an artifact-free FD-OCT depth profile without the requirement of complicated equipment. The proposed Pix2Pix GAN–based FD-OCT system has a simple design and relatively low cost.

The remainder of this paper is organized as follows. Section II describes the conventional FD-OCT architecture and artifact problems in the FD-OCT system along with image-to-image translation by using the Pix2Pix GAN. Section III presents the experimental results and a discussion of them. Finally, Section IV provides concluding remarks and suggestions for future studies.

II. PIX2PIX GAN–BASED FULL-RANGE FD-OCT

In this study, we propose a simplified FD-OCT scheme by using Pix2Pix GAN–based image-to-image translation to remove unwanted artifacts.

A. CONVENTIONAL FULL-RANGE FD-OCT

FD-OCT systems are based on the measurement of the interference spectrum, either in space on a spectrometer or in time during the wavelength sweep of a rapidly tunable laser source. Figure 1 depicts the architecture of the conventional full-range FD-OCT system implemented in this study to create the dataset. This system consists of a super luminescent diode (SLD), a 2 × 2 coupler, two collimators, a reference mirror, and an optical spectrum analyzer (OSA) that are used to obtain OCT spectra. The SLD, a broadband low-coherence light source, is distributed to the sample and reference arm through the coupler. In the sample arm, the beam is directly delivered to the sample, and then scattered and reflected into a collimator. In the reference arm, the probing beam is incident on a mirror. Let parameter n_1 represents the refractive index of the sample, d_0 indicates the front surface of the sample toward the zero-delay position, and d_1 denotes the physical thickness of the sample. Light emerging from the mirror passes back to the coupler in the reference arm. Beams from the reference and sample arms are recombined in the coupler and then passed to the OSA. Subsequently, the received spectra use the IFFT method for signal processing.

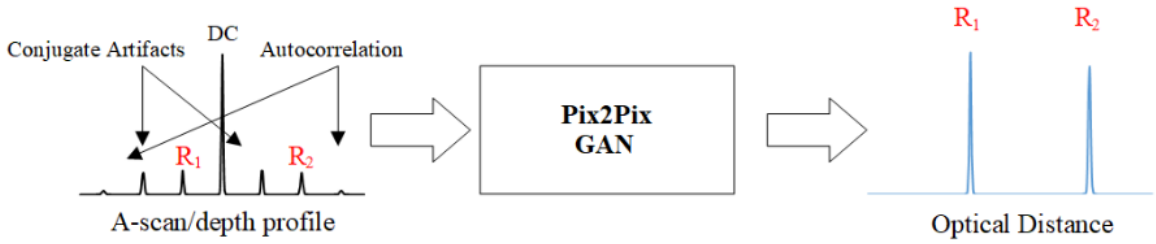

FIGURE 2. Pix2Pix GAN-based FD-OCT system.

The inverse Fourier transform of the intensity signal is given in [24]:

$$\begin{aligned} I_D(k) &= \frac{\rho}{8}[S(k)(R_R + R_{S1} + R_{S2} + \ldots)] \\ &+ \frac{\rho}{4}\left[\sum_{n=1}^{N}\sqrt{R_R R_{Sn}}\{S[2(z_R - z_{Sn})]\}\right. \\ &+ \left.\{S[-2(z_R - z_{Sn})]\}\right] \\ &+ \frac{\rho}{8}\left[\sum_{n\neq m}^{N}\sqrt{R_{Sm}R_{Sn}}\{S[2(z_{Sm} - z_{Sn})]\}\right. \\ &+ \left.\{S[-2(z_{Sm} - z_{Sn})]\}\right] \end{aligned} \quad (1)$$

where k is the wavenumber, $I_D(k)$ is the photocurrent, ρ is the responsivity of the detector (A/W), $S(k)$ is the power spectral, z_R indicates the distance from the beam splitter or fiber coupler to the reference reflector, z_s represents the path length variable in the sample arm measured from the beam splitter, and R_R and R_S denote the power reflectivity of the reference reflector and each reflector in the sample arm, respectively. From (1), the photocurrent can be divided into three components. The first term is called the DC component, and its amplitude is proportional to the power reflectivity of the reference mirror added to the sum of sample reflectivity. The DC component has the largest amplitude of the detector current and is an artifact in OCT. The second item is the desired component of OCT, which is the cross-correlational component of each sample reflector. The last component is an autocorrelation part, which is also an artifact in OCT, caused by the interference between different sample reflectors. In addition, a conjugate artifact, also called a mirror artifact, is another type of artifact found in OCT. This artifact occurs because the detected interferometric spectrum is real, whereas inverse Fourier transform of the spectral shows Hermitian symmetry. In other words, it's positive and negative distances are the complex conjugates of each other. Therefore, if they are real, they must be identical [24].

In OCT systems, the axial resolution is determined by the coherence length of the light source and defined using the Rayleigh criterion. The axial resolution of the OCT system can be expressed as follows:

$$\Delta z = \frac{l_c}{2} = 0.44\frac{\lambda_0^2}{\Delta\lambda} \quad (2)$$

where l_c is the coherence length of the light source, λ_0 is the central wavelength of the light source, and $\Delta\lambda$ is the bandwidth.

The maximal depth position Z_{max} was determined according to the OSA wavelength resolution and is given as follows:

$$Z_{max} = \frac{1}{4}.\frac{\lambda_0^2}{\delta\lambda} \quad (3)$$

where $\delta\lambda$ is the OSA wavelength resolution.

Assume that the central wavelength of the light source is 1550 nm and the full width at half maximum of the spectrum is 20 nm. By using (2), the value of axial resolution can be calculated as approximately 52.86 μm.

B. PIX2PIX GAN–BASED FD-OCT

To eliminate these artifacts by using simple and cost-effective methods, we propose applying the Pix2Pix GAN to the FD-OCT system to obtain an artifact-free FD-OCT depth profile, as shown in Figure 2. The Pix2Pix GAN is a general approach for image-to-image translation. It is a type of a conditional GAN, where the generation of the output image is conditional on the input image [25]. Compared with other GAN models, the conditional GAN has the capability of generating large high-quality images for a variety of image translation tasks. Therefore, the Pix2Pix GAN has been widely used to train a deep convolutional neural network and generates data that is similar to real data.

TABLE 1. Devices comparison.

Devices	Method	
	Conventional Full-Range FD-OCT	Pix2Pix GAN-Based FD-OCT
Hardware	Polarization Plate	Standard Computer
	Beam Splitter	GPU with 11 GB of DRAM
	Quarter-Wave Plate	
	Mirror	
	Collimator	
Software	-	Open-Source Python
		Open-Source TensorFlow

Table 1 presents the devices required for the conventional FD-OCT method proposed in our previous work [5] and the

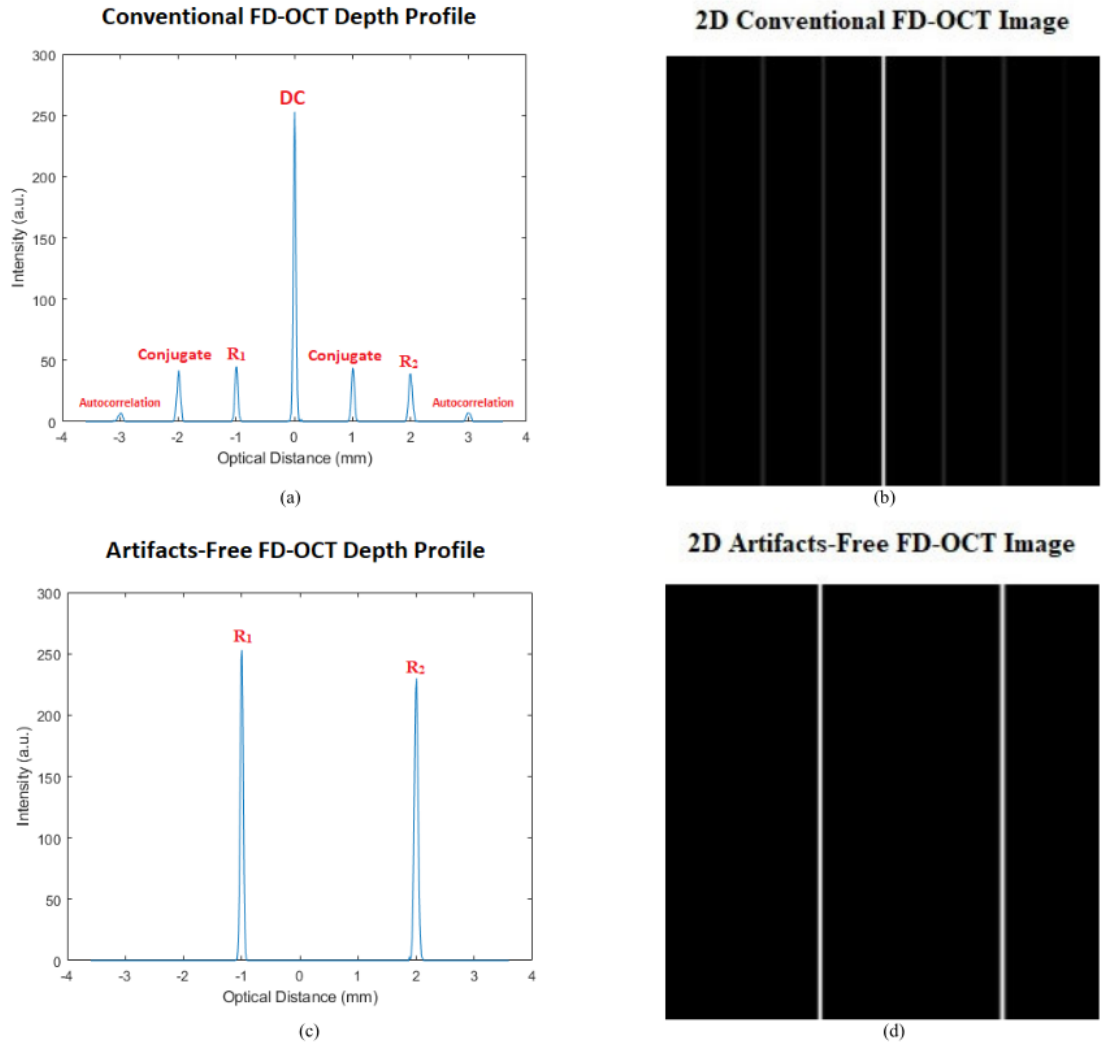

FIGURE 3. Sample of the FD-OCT dataset. (a) Conventional FD-OCT depth profile (b) 2D conventional FD-OCT image (c) Artifact-free FD-OCT depth profile (d) 2D artifact-free FD-OCT image.

proposed Pix2Pix GAN-based FD-OCT. The conventional FD-OCT method consist of two polarization plates, one beam splitter, one quarter-wave plate, two mirrors, and two collimators. Meanwhile, the proposed Pix2Pix GAN-based FD-OCT is computationally inexpensive and can be run on standard computers without the need of complex and expensive optical equipment, verified the cost-effective feature of the proposed system. The software used in the proposed system is the open-source framework.

The dataset used for training the Pix2Pix GAN was generated by implementing conventional full-range FD-OCT and contained many pairs of images. Each pair of images included an original image and its expected transformed result. In other words, we used the Pix2Pix GAN to translate a conventional FD-OCT depth profile into an artifact-free FD-OCT depth profile. First, the conventional FD-OCT depth profiles were generated using (1) with different parameters. Subsequently, the corresponding artifact-free FD-OCT depth profiles were obtained using the phase-shifting algorithm proposed in our previous study [4]. However, as shown in Figure 3 (a) and (c), these conventional and artifact-free FD-OCT depth profiles were all contained one-dimensional (1D) data, which could not be directly fed to the Pix2Pix GAN. To solve this problem, without the loss of generality, we expanded these 1D FD-OCT depth profiles into two-dimensional (2D) images by duplicating original 1D data. Assume that the size of the conventional (or artifact-free) FD-OCT depth profile is $1 \times N$, where N denotes sampling points. Subsequently, raw data

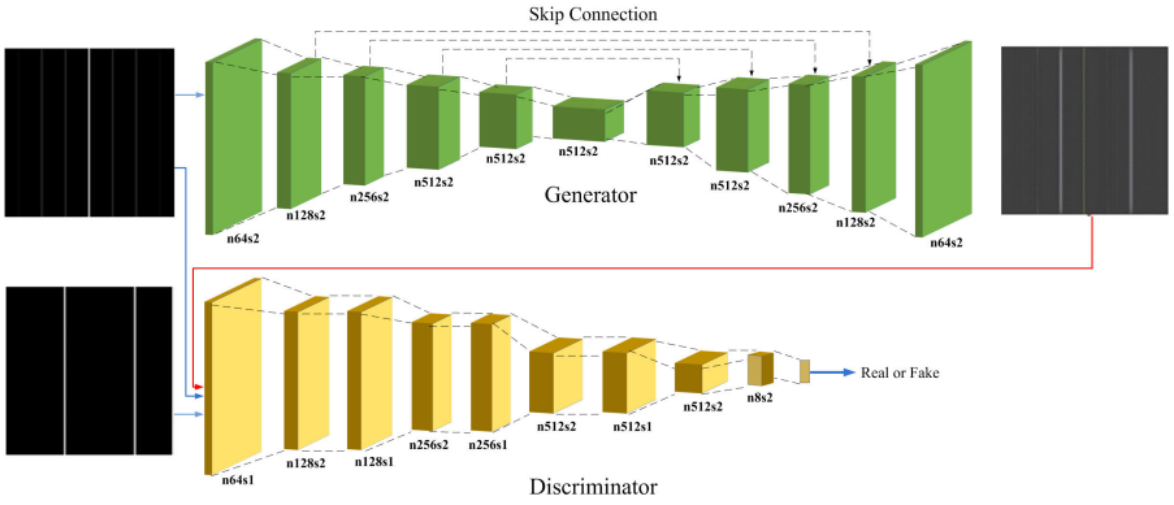

FIGURE 4. Pix2Pix GAN architecture.

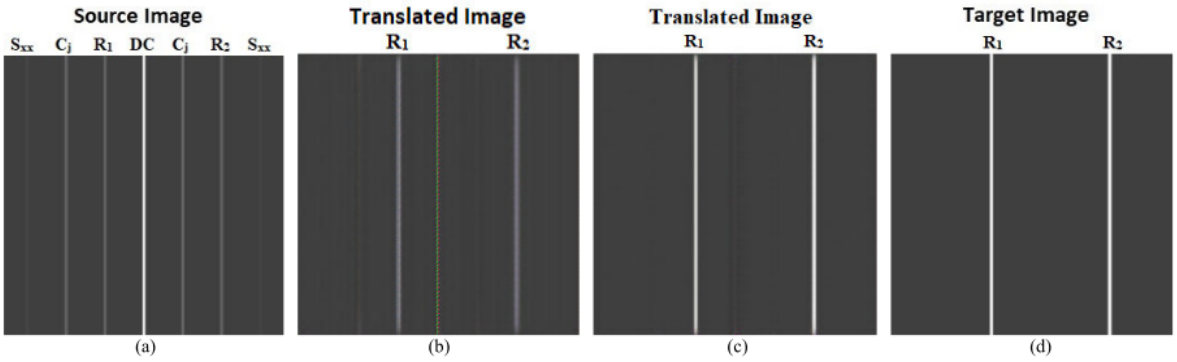

FIGURE 5. Image translation results for Pix2Pix GAN–based FD-OCT.

were copied and expanded to an $N \times N$ FD-OCT image (i.e., Figure 3 (b) and (d)). In addition, the brightness of lines inside the corresponding 2D image was proportional to the intensity of the FD-OCT depth profile.

Figure 4 shows the Pix2Pix GAN architecture used in this study where n denotes the batch size and s is the stride. The discriminator was based on a PatchGAN model, which classified patches of an image as real or fake by outputting a matrix of values as the output instead of a single value. The advantage of using this model was that it provided sharp high-frequency details and the number of parameters could also be reduced. Source and target images (i.e., the 2D conventional and artifact-free FD-OCT images, respectively) were provided to the discriminator to determine whether the target was real or fake.

The generator had a U-shaped network architecture, which added skipped connections between each central symmetric layer to prevent the loss of small information. The generator was used an encoder–decoder network composed of the standardized blocks of convolutional, batch normalization, dropout, and activation layers. The generator was trained through adversarial loss and updated through $L1$ loss that was measured between the generated image and the expected output image. This additional loss drove the generator model to create plausible translations of the source image. Overall, the generator was updated through a weighted sum of both adversarial and $L1$ losses. To achieve this simultaneous training, the logical or composite model was used to stack the generator on top of the discriminator. A source image was provided as the input to the generator and discriminator. In addition, the output of the generator was provided to the discriminator as the plausible image. The discriminator then predicted the likelihood of an image being real or fake, thus translating the source image. The objective of the proposed model is expressed as follows [19]:

$$L_{GAN}(G, D) = E_{x,y}\left[\log D(x, y)\right] + E_x\left[\log\left(1 - D(x, G(x))\right)\right] \quad (4)$$

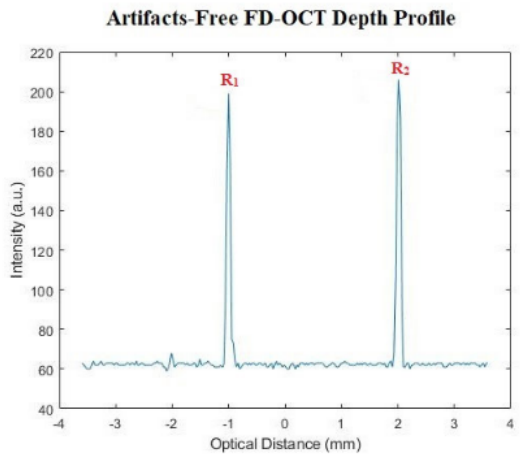

FIGURE 6. The 1D artifact-free FD-OCT depth profile after 100 epochs.

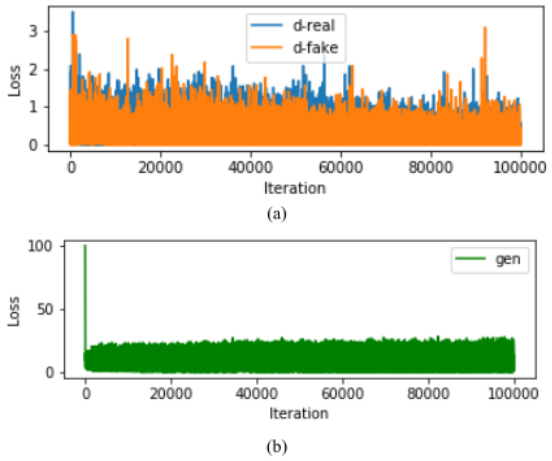

FIGURE 7. Loss of GAN-based FD-OCT system. (a) Discriminator loss (b). Generator loss.

where x is the source image, and y is the corresponding target image, set as the ground truth for x. In (4), the generator G attempted to minimize this objective in response to an adversarial discriminator D that attempted to maximize it, where the results are optimized as follows:

$$G^* = \arg\min_G \max_D L_{GAN}(G, D) \tag{5}$$

where G^* is the resulting optimized generator.

The $L1$ loss from the generated and expected output images can be evaluated as follows:

$$L1(G) = E_{x,y}\left[\|y - G(x)\|_1\right] \tag{6}$$

The final objective of the proposed model can be denoted as follows:

$$G^* = \arg\min_G \max_D L_{GAN}(G, D) + \lambda L1(G) \tag{7}$$

TABLE 2. Optical distance of FD-OCT.

	Optical Distance (mm)	
	R_1	R_2
Translated Image	-1.0007	2.0155
Target Image	-1.0007	2.0155

TABLE 3. FD-OCT image parameter.

	Type	
	A、C	B、D
n_1	1.5	2.5
d_0	-500μm	-500μm
d_1	2,000μm	1,200μm

To train the discriminator, batches of real and fake images are required. The corresponding discriminator of the real image is given the label of class = 1 to indicate that they are real, whereas the corresponding discriminator of the fake image is given the label of class = 0 to indicate that they are fake.

In this experiment, we used the 2D FD-OCT image dataset for both training and validation. The objective of image translation was to convert 2D conventional FD-OCT images into 2D artifact-free FD-OCT images. Both the training and validation datasets contained 1,000 images each. Both 2D conventional FD-OCT images and 2D artifact-free FD-OCT images were in JPEG format with an image being 600 pixels wide and 600 pixels high.

The model was trained using the Keras DL framework on a personal workstation with an NVIDIA GeForce RTX 2080 Ti graphics processing unit with 11 GB of memory. Each image was loaded and paired between a 2D conventional FD-OCT image and a 2D artifact-free FD-OCT image. For computational purposes, the datasets were downscaled to images that were 256 pixels wide and 256 pixels high. The arrays of the datasets were saved in compressed NumPy array format. The discriminator model was optimized using the Adam optimizer with a learning rate of 0.0002 and a beta of 0.5 and by implementing the binary cross entropy loss function. The generator model applied different activation functions for the encoder and decoder network. Because small (negative) value will be induced in the down-sampling model, Leaky ReLU was used in the encoder network (down-sampling) to prevent the "dead ReLU" problem. However, for the decoder network (up-sampling), ReLU activation function still work well and have the advantages of faster performance and introduce more non-linearity [17], [26]. For the output layer, tanh activation function was implemented to normalized the output in the range of [-1,1] since the range of the brightness/value of each pixel of real image should be within [-1,1]. For weight initialization, a random normal with a standard deviation of 0.02 was applied [27].

GAN models typically do not converge; therefore, an equilibrium was found between the generator and discriminator

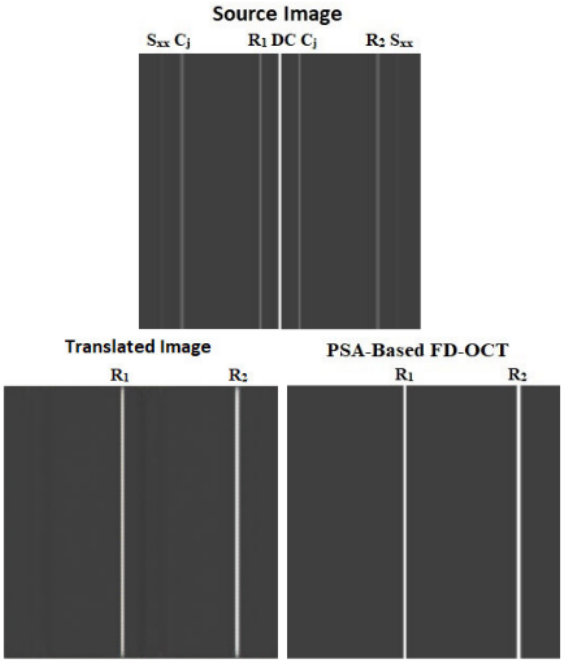

FIGURE 8. The target image and translated results for 2D conventional FD-OCT image of type A source image.

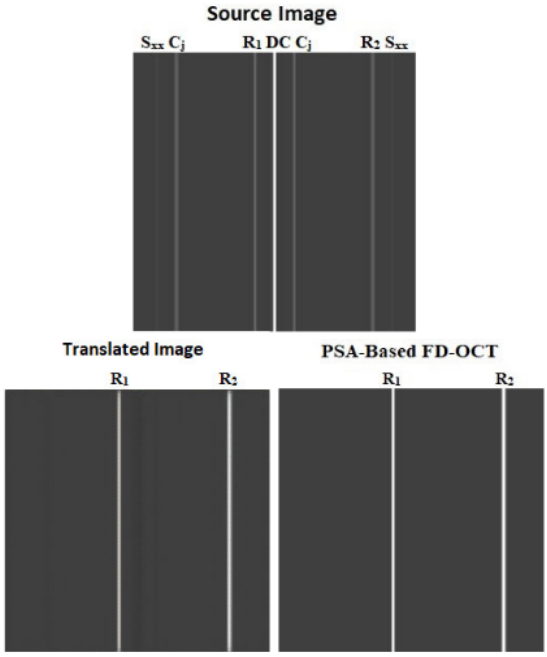

FIGURE 9. The target image and translated results for 2D conventional FD-OCT image of type B source image.

models; that is, the image quality was used to choose the best model. To examine the quality of the translated image, the model was saved to an *H5* formatted file every 10 training epochs and used later to generate image-to-image translations. The total number of epochs was set to 100. The results may have varied due to the stochastic nature of the algorithm in terms of differences in the numerical precision.

RESULTS AND DISCUSSION

In this section, we present the experimental results of the proposed Pix2Pix GAN-based FD-OCT model. The model was trained with training dataset of 1,000 images. Test images were chosen randomly from the validation dataset with 1,000 images. The results were observed every 10 epochs by using the saved model to generate translated images. Three experiments were conducted with the trained model.

First experiment aims to prove the feasibility of the Pix2Pix GAN-based FD-OCT. Figure 5 displays the source, translated, and target images (i.e., the ground truth image) of FD-OCT, respectively. The indices of S_{xx}, C_j, DC, R_1 and R_2 indicate the auto-correlation artifact, conjugate artifact, DC artifact, first and second peaks of FD-OCT, respectively. The source, translated, and target images denote the 2D conventional FD-OCT image, 2D artifact-free FD-OCT image produced using the Pix2Pix GAN, and real 2D artifact-free FD-OCT image, respectively. Our purpose was to translate 2D conventional FD-OCT images into 2D artifact-free FD-OCT images. From the translated images, we observed that the proposed model could translate the input FD-OCT image into the desired result. The FD-OCT depth profile of the translated image was the same as that of the target image. Figure 5 (b) reveals that some background noise was still present after 10 training epochs. However, after 100 training epochs, the translated image was clearer and approaching the target image, as shown in Figure 5 (c).

To verify the optical distance of the FD-OCT image, we reduced the dimension of the translated image to obtain the corresponding 1D FD-OCT depth profile, as shown in Figure 6. The first and second peaks of the FD-OCT depth profile (R_1 and R_2) obtained from the translated image matched those of the FD-OCT depth profile obtained from the target image (i.e., -1.0007 and 2.0155, respectively). Note that since the Pix2Pix-GAN generates image in RGB mode, the black color of the translated image is not truly black, cause an average bias about 60 in the intensity. Table 2 lists the optical distances of R_1 and R_2 for both translated and target images.

The performance of the discriminator and generator models was recorded for each iteration. Figure 7 shows the discriminator loss for real images (blue), discriminator loss for generated fake images (orange), and the generator loss for generated fake images (green) for 100,000 iterations. Discriminator loss for real and fake samples is about the same at around 1.0 and loss for the generator is higher. This result

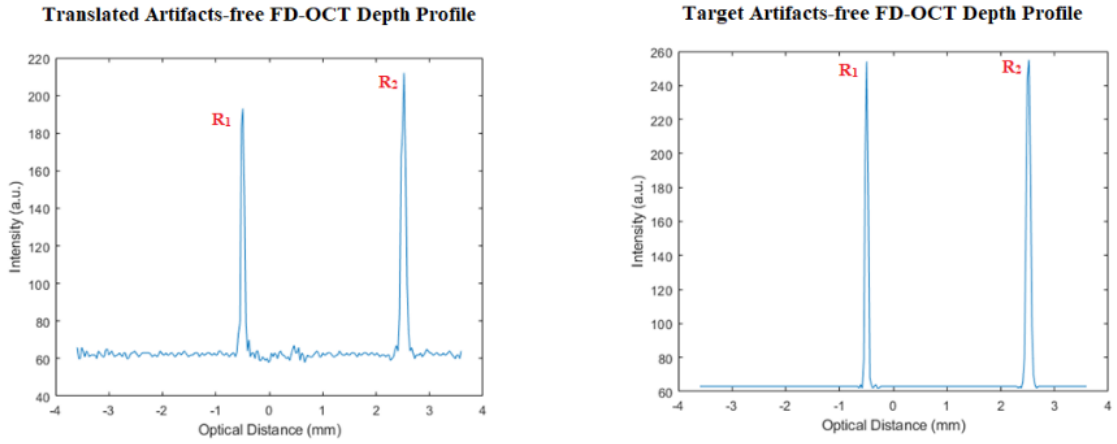

FIGURE 10. The 1D translated and target artifacts-free FD-OCT depth profiles of type A image.

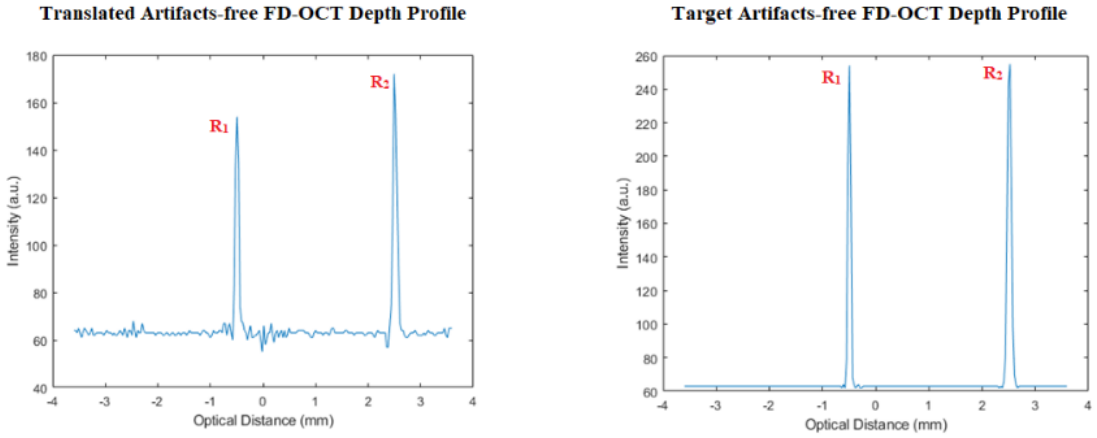

FIGURE 11. The 1D translated and target artifacts-free FD-OCT depth profiles of type B image.

indicates the stable GAN with normal loss during the training process.

Further experiments are conducted without re-train the network and used the same weight obtained from previous training. The objective of image translation is to convert 2D conventional FD-OCT images into 2D artifact-free FD-OCT images. Four types of images are constructed by using the hyper-parameters in Table 3 and their specifications are described as follows:

1. Type A: one of the 2D conventional FD-OCT images from the validation dataset.
2. Type B: a 2D conventional FD-OCT image, which possess the same d_0 with type A, whereas n_1 and d_1 are different.
3. Type C: the hyper-parameters are the same as type A, but random Gaussian noise with SNR of 0 (or 3 dB) is added on the power spectra form OSA.
4. Type D: the image is constructed by adding random Gaussian noise on the power spectra from OSA with SNR of 0 dB (or 3 dB) to type B image.

Note that all these four types of images are not included in the training dataset.

Second experiment applied the images of types A and B to verify the capability of the proposed Pix2Pix GAN-based FD-OCT model in reconstructing 2D artifact-free FD-OCT images of unknown 2D conventional FD-OCT images. Note that we also utilized the phase-shift algorithm (PSA) in [4] to reconstruct these test images for comparison. Since these images are not included in the training dataset, to avoid confusion, these images are names as PSA-based FD-OCT images in the rest of this study.

Figures 8 and 9 show the target image and translated results for 2D conventional FD-OCT image of types A and B, respectively. It has been shown that the model could translate the

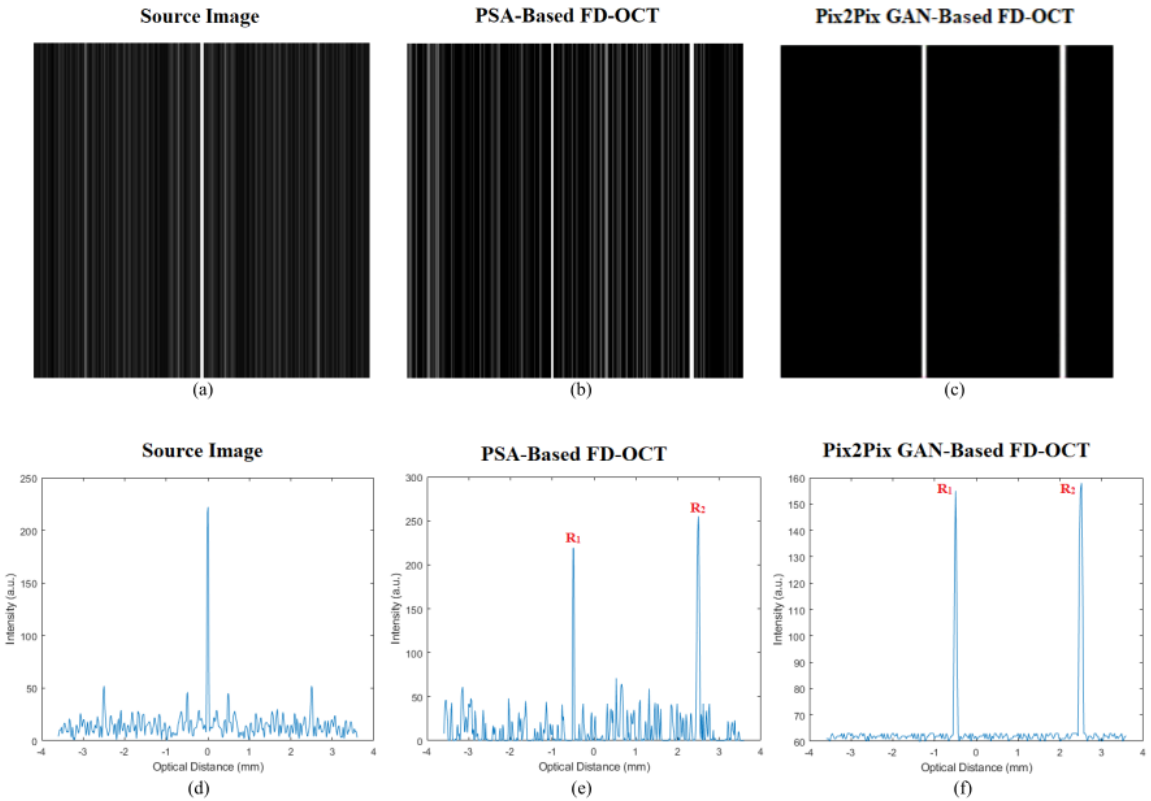

FIGURE 12. The 1D FD-OCT depth profiles and 2D images of the type C source images of SNR 0 dB, PSA-based FD-OCT image, and Pix2Pix GAN-based FD-OCT image.

source image into a 2D artifact-free FD-OCT image, which is the same as the PSA-based FD-OCT image. Moreover, to confirm the optical distance of the translated image, these 2D images are converted to the 1D FD-OCT depth profiles by reducing the dimension. Figures 10 and 11 show the 1D translated and target artifacts-free FD-OCT depth profiles of types A and B, respectively.

For type A, the first peak (R_1) and second peak (R_2) of the FD-OCT depth profile from the translated image matched those of the target image (-0.4933 and 2.5228), respectively. Meanwhile, for type B, the first peak (R_1) and second peak (R_2) of the FD-OCT depth profile from the translated image matched those of the target image, (-0.4933 and 2.4947), respectively. The second peak represented a small error but was in the acceptable range. The optical distance of R_1 and R_2 both for translated and target images are shown in Table 4.

In order to prove that the proposed Pix2Pix GAN-based FD-OCT can eliminate the artifacts of FD-OCT image with more complex noisy, the third experiment used the images from the validation images of types C and D.

Figure 12 presents the 1D FD-OCT depth profiles and 2D images of the type C source image of SNR 0 dB, PSA-based FD-OCT image, and Pix2Pix GAN-based FD-OCT image, respectively. As shown in Figure 12 (a) and (d), it can be observed that there are a lot of noises in the source image, except the artifacts. Figure 12 (b) and (e) depict the results of PSA-based FD-OCT, where noises still exist expect for the peaks of FD-OCT. Further, Figure 12 (c) and (f) show the translated image using the Pix2Pix GAN-based FD-OCT method. It has been shown that the artifacts and noises are almost eliminated, and only an acceptable small noise-floor remains in the image. Figure 13 depicts the 1D FD-OCT depth profiles and 2D images of the type C source image of SNR 3 dB, PSA-based FD-OCT image, and Pix2Pix GAN-based FD-OCT image, respectively.

TABLE 4. FD-OCT optical distance comparison.

	Optical Distance (mm)	
	R_1	R_2
Type A		
Translated Image	-0.4933	2.5228
Target Image	-0.4933	2.5228
Type B		
Translated Image	-0.4933	2.5228
Target Image	-0.4933	2.4947

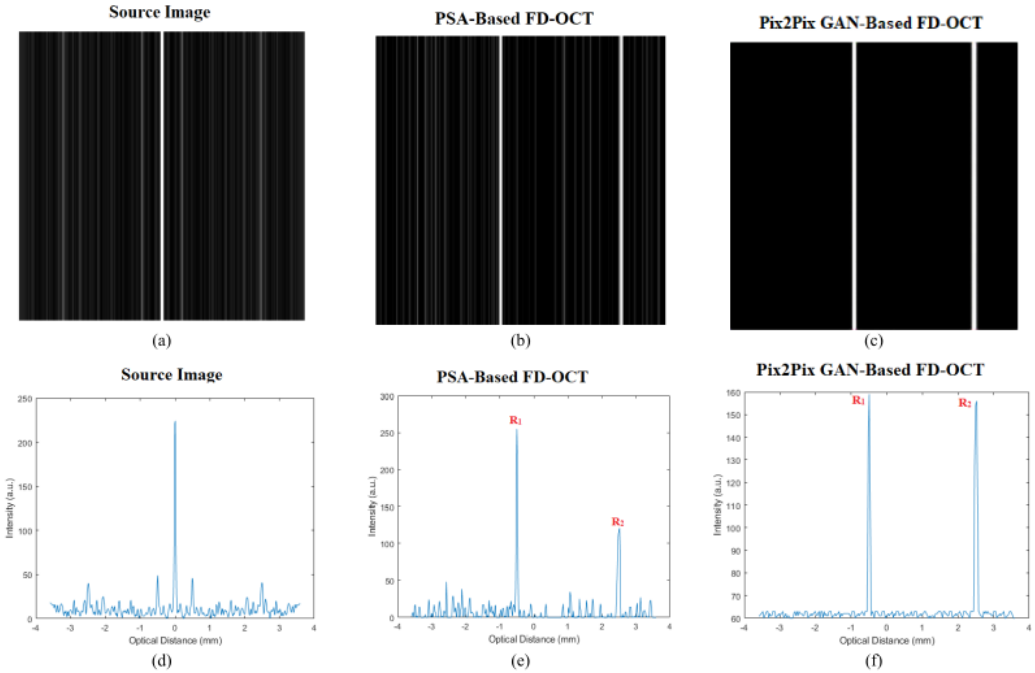

FIGURE 13. The 1D FD-OCT depth profiles and 2D images of the type C source images of SNR 3 dB, PSA-based FD-OCT image, and Pix2Pix GAN-based FD-OCT image.

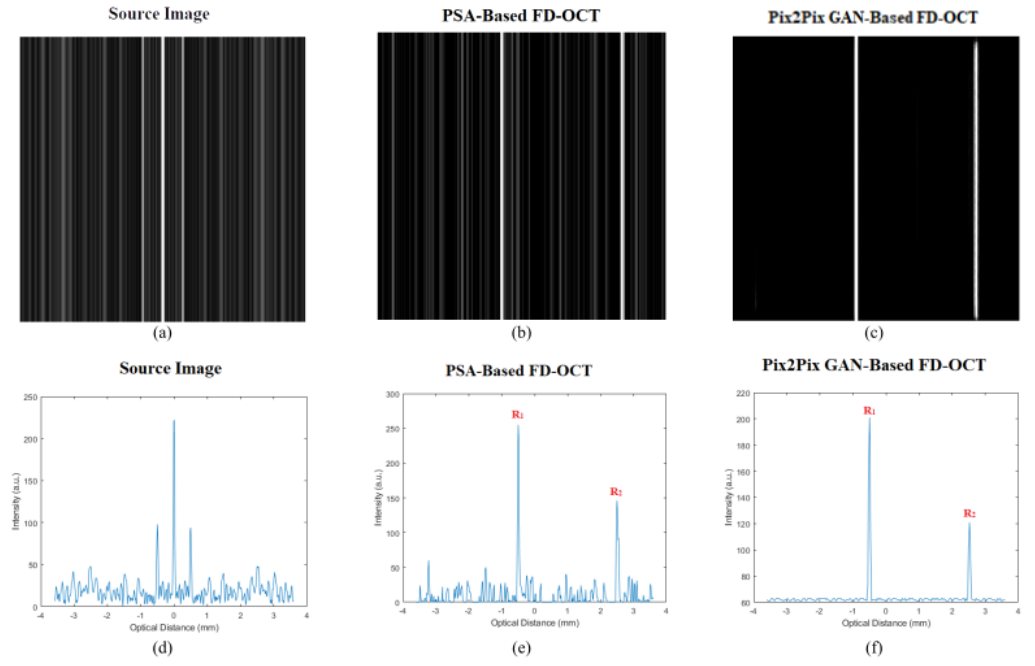

FIGURE 14. The 1D FD-OCT depth profiles and 2D images of the type D source images of SNR 0 dB, PSA-based FD-OCT image, and Pix2Pix GAN-based FD-OCT image.

From Figure 13 (a) and (d), we can observe that much noises appear in the source image. Figure 13 (b) and (e) display the results of PSA-based FD-OCT, which still have noises in addition to the FD-OCT peaks. Figure 13 (c) and (f) present

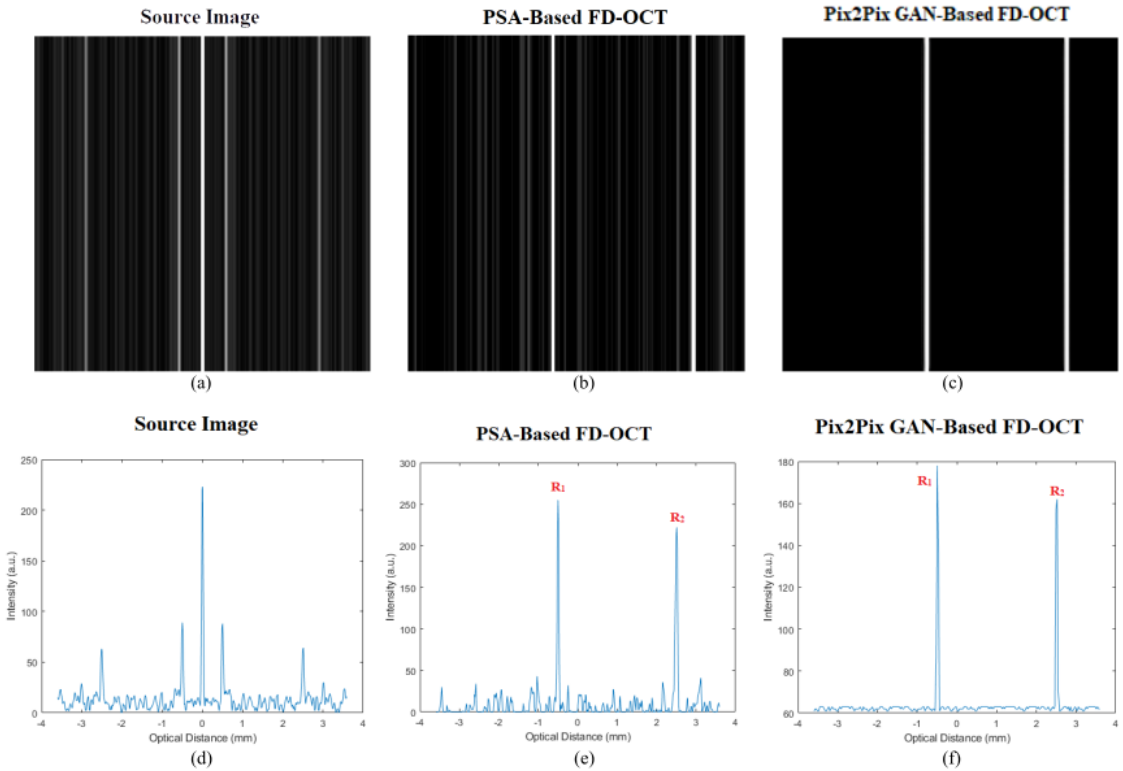

FIGURE 15. The 1D FD-OCT depth profiles and 2D images of the type D source images of SNR 3 dB, PSA-based FD-OCT image, and Pix2Pix GAN-based FD-OCT image.

TABLE 5. FD-OCT optical distance comparison.

	Optical Distance (mm)	
	R_1	R_2
Type C 0 dB SNR		
Translated Image	-0.4933	2.5228
Target Image	-0.4933	2.5228
Type C 3 dB SNR		
Translated Image	-0.4933	2.5228
Target Image	-0.4933	2.5228

TABLE 6. FD-OCT optical distance comparison.

	Optical Distance (mm)	
	R_1	R_2
Type D 0 dB SNR		
Translated Image	-0.4933	2.5228
Target Image	-0.4933	2.5228
Type D 3 dB SNR		
Translated Image	-0.4933	2.5228
Target Image	-0.4933	2.5228

the translated image using the Pix2Pix GAN-based FD-OCT method, where artifacts are removed and only a low acceptable noise-floor remains. These results verify that the proposed Pix2Pix GAN-based FD-OCT can effectively remove the artifacts and other noises in the resulted artifacts-free FD-OCT image. Compare to the PSA in [4], the proposed Pix2Pix GAN-based FD-OCT achieves superior results.

Table 5 presents the optical distance of R_1 and R_2 for type C source image of SNR 0 dB (or 3 dB) reconstructing by using PSA-based FD-OCT and Pix2Pix GAN-based FD-OCT method, respectively. It can be observed that the first peak (R_1) and second peak (R_2) of the FD-OCT depth profile reconstructed by the Pix2Pix GAN-based FD-OCT are the same as that of PSA-based FD-OCT.

Figures 14 and 15 depict the 1D FD-OCT depth profiles and 2D images of the type D source images of SNR 0 dB and 3 dB, PSA-based FD-OCT image, and Pix2Pix GAN-based FD-OCT image, respectively. It has been shown that a lot of noises are accompanied with the source image. The PSA-based FD-OCT can remove the artifacts, but still much noises exist except for the FD-OCT peaks. In addition, the proposed Pix2Pix GAN-based FD-OCT can eliminate the artifacts and effectively reduce the noise power. Similarly, the optical distance of R_1 and R_2 for type D

source image of SNR 0 dB and 3 dB are respectively reconstructed by using PSA-based FD-OCT and Pix2Pix GAN-based FD-OCT method, as shown in Table 6. The values of R_1 and R_2 of the FD-OCT depth profile reconstructed from the Pix2Pix GAN-based FD-OCT and the PSA-based FD-OCT are the same.

Table 6 shows the optical distance of R_1 and R_2 from type D both for translated image from Pix2Pix GAN and target images. For type D with 0 dB SNR, the first peak (R_1) and second peak (R_2) of the FD-OCT depth profile from the translated image matched those of the target image (-0.4933 and 2.5228), respectively. Furthermore, for type D with 3 dB SNR, the first peak (R_1) and second peak (R_2) of the FD-OCT depth profile from the translated image matched those of the target image, (-0.4933 and 2.5228), respectively.

IV. CONCLUSION

In this study, one Pix2Pix GAN-based FD-OCT is proposed to eliminate the artifacts of FD-OCT images, including conjugate, DC, and auto-correlation artifact. Four types of 2D conventional FD-OCT images (i.e., types A to D) are constructed to verify the reconstructing capability of model. The first two experiment results show that the proposed Pix2Pix GAN-based FD-OCT can reconstruct the 1D FD-OCT depth profile and 2D artifact-free FD-OCT images as well as the phase-shift algorithm-based FD-OCT system. Further, at the last experiment, additional Gaussian noise of SNR 0 dB and 3 dB are respectively added to the power spectral to generate noisier 2D conventional FD-OCT images (i.e., types C and D) for testing the model. From the results, it can be observed that the proposed model can effectively remove the artifacts and other noises in the resulted artifacts-free FD-OCT image. In addition, the proposed Pix2Pix GAN–based FD-OCT has advantages of simple design and relatively low cost. Our future work will focus on how to apply this model to more complex FD-OCT depth profile structure.

REFERENCES

[1] S. Asrani, L. Essaid, B. D. Alder, and C. S. Turla, "Artifacts in spectral-domain optical coherence tomography measurements in glaucoma," *JAMA Ophthalmol*, vol. 132, no. 4, pp. 396–402, 2014, doi: 10.1001/jamaophthalmol.2013.7974.

[2] J. Chhablani, T. Krishnan, V. Sethi, and I. Kozak, "Artifacts in optical coherence tomography," *Saudi J. Ophthalmology*, vol. 28, no. 2, pp. 81–87, Apr. 2014, doi: 10.1016/j.sjopt.2014.02.010.

[3] J. Lin, S. Zhong, Q. Zhang, and W. Chen, "Five-frame variable phase-shifting method for full-range spectral-domain optical coherence tomography," *Appl. Sci.*, vol. 8, no. 9, p. 1580, Sep. 2018, doi: 10.3390/app8091580.

[4] H.-C. Cheng, J.-F. Huang, and Y.-H. Hsieh, "Numerical analysis of one-shot full-range FD-OCT system based on orthogonally polarized light," *Opt. Commun.*, vol. 282, no. 14, pp. 3040–3045, Jul. 2009, doi: 10.1016/j.optcom.2009.04.024.

[5] H. C. Cheng and M. S. Shiu, "Experimental demonstration of high-speed full-range Fourier domain optical coherence tomography imaging using orthogonally polarized light and a phase-shifting algorithm," *Appl. Opt.*, vol. 51, pp. 8762–8768, Dec. 2012, doi: 10.1364/AO.51.008762.

[6] W. Bao, Y. Shen, T. Chen, P. Li, and Z. Ding, "High-speed high-precision and ultralong-range complex spectral domain dimensional metrology," *Opt. Exp.*, vol. 23, no. 9, May 2015, Art. no. 11013, doi: 10.1364/OE.23.011013.

[7] Q. Zhang, S. Zhong, and J. Zhong, "Sine-modulated wavelength-independent full-range complex spectral optical coherence tomography with an ultra-broadband light source," *Adv. Mech. Eng.*, vol. 7, no. 5, May 2015, Art. no. 168781401558872, doi: 10.1177/1687814015588726.

[8] G. J. Liu, O. Tan, S. S. Gao, A. D. Pechauer, B. K. Lee, C. D. Lu, J. G. Fujimoto, and D. Huang, "Postprocessing algorithms to minimize fixed-pattern artifact and reduce trigger jitter in swept source optical coherence tomography," *Opt. Exp.*, vol. 23, pp. 9824–9834, Apr. 2015, doi: 10.1364/OE.23.009824.

[9] J. Lee, T. Yoon, and B. H. Lee, "Post-processing method for image reconstruction enhancement in integrating-bucket-based full-field optical coherence tomography," *Appl. Sci.*, vol. 10, no. 3, p. 830, Jan. 2020, doi: 10.3390/app10030830.

[10] Y. Zhu and W. R. Gao, "Single-shot wavelength-independent phase-shifting method for full-field optical coherence tomography," *Appl. Opt.*, vol. 58, p. 806, 2019, doi: 10.1364/AO.58.000806.

[11] X. Y. Liu, M. Y. Ke, X. W. Yao, J. Chua, L. Schmetterer, and B. Y. Tan, "Stable complex conjugate artifact removal in OCT using circularly polarized light as reference," *Opt. Lett.*, vol. 45, pp. 3977–3980, Jul. 2020, doi: 10.1364/OL.395860.

[12] H.-C. Cheng and C.-T. Huang, "Measurement of thickness and refractive index of optical samples simultaneously using full-range one-shot frequency-domain optical coherence tomography," *Fiber Integr. Opt.*, vol. 34, no. 3, pp. 145–156, May 2015, doi: 10.1080/01468030.2015.1044678.

[13] S. K. Devalla, G. Subramanian, T. H. Pham, X. Wang, S. Perera, T. A. Tun, T. Aung, L. Schmetterer, A. H. Thiéry, and M. J. A. Girard, "A deep learning approach to denoise optical coherence tomography images of the optic nerve head," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, Art. no. 14454, doi: 10.1038/s41598-019-51062-7.

[14] B. Buchroithner, A. Prylepa, P. J. Wagner, S. E. Schausberger, D. Stifter, and B. Heise, "Full-field optical coherence tomography in a balanced detection mode," *Appl. Opt.*, vol. 57, pp. 8705–8710, Oct. 2018, doi: 10.1364/AO.57.008705.

[15] H. Cheong, S. K. Devalla, T. H. Pham, L. Zhang, T. A. Tun, X. Wang, S. Perera, L. Schmetterer, T. Aung, C. Boote, A. Thiery, and M. J. A. Girard, "DeshadowGAN: A deep learning approach to remove shadows from optical coherence tomography images," *Transl. Vis. Sci. Technol.*, vol. 9, no. 2, p. 23, Apr. 2020, doi: 10.1167/tvst.9.2.23.

[16] C.-H. Lin, W.-M. Liao, J.-W. Liang, P.-H. Chen, C.-E. Ko, C.-H. Yang, and C.-K. Lu, "Denoising performance evaluation of automated age-related macular degeneration detection on optical coherence tomography images," *IEEE Sensors J.*, vol. 21, no. 1, pp. 790–801, Jan. 2021, doi: 10.1109/JSEN.2020.3014254.

[17] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014, *arXiv:1406.2661*. [Online]. Available: http://arxiv.org/abs/1406.2661

[18] H. Cheong, S. K. Devalla, T. Chuangsuwanich, T. A. Tun, X. Wang, T. Aung, L. Schmetterer, M. L. Buist, C. Boote, A. H. Thiéry, and M. J. A. Girard, "OCT-GAN: Single step shadow and noise removal from optical coherence tomography images of the human optic nerve head," 2020, *arXiv:2010.11698*. [Online]. Available: http://arxiv.org/abs/2010.11698

[19] Q. Hao, K. Zhou, J. Yang, Y. Hu, Z. Chai, Y. Ma, G. Liu, Y. Zhao, S. Gao, and J. Liu, "High signal-to-noise ratio reconstruction of low bit-depth optical coherence tomography using deep learning," *J. Biomed. Opt.*, vol. 25, no. 12, Nov. 2020, Art. no. 123702, doi: 10.1117/1.JBO.25.12.123702.

[20] A. Guo, L. Fang, M. Qi, and S. Li, "Unsupervised denoising of optical coherence tomography images with nonlocal-generative adversarial network," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021, doi: 10.1109/TIM.2020.3017036.

[21] Y. Huang, W. Xia, Z. Lu, Y. Liu, H. Chen, J. Zhou, L. Fang, and Y. Zhang, "Noise-powered disentangled representation for unsupervised speckle reduction of optical coherence tomography images," *IEEE Trans. Med. Imag.*, early access, Dec. 16, 2020, doi: 10.1109/TMI.2020.3045207.

[22] T. Zhang, J. Cheng, H. Fu, Z. Gu, Y. Xiao, K. Zhou, S. Gao, R. Zheng, and J. Liu, "Noise adaptation generative adversarial network for medical image analysis," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1149–1159, Apr. 2020, doi: 10.1109/TMI.2019.2944488.

[23] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976, doi: 10.1109/CVPR.2017.632.

[24] J. A. Izatt and M. A. Choma, "Theory of optical coherence tomography," in *Optical Coherence Tomography-Technology and Applications*. Berlin, Germany: Springer, 2015, pp. 47–72. [Online]. Available: https://www.springer.com/gp/book/9783319064185

[25] T. E. Tavolara, M. K. K. Niazi, V. Arole, W. Chen, W. Frankel, and M. N. Gurcan, "A modular cGAN classification framework: Application to colorectal tumor detection," *Sci. Rep.*, vol. 9, no. 1, Art. no. 18969, Dec. 2019, doi: 10.1038/s41598-019-55257-w.

[26] M. Mehralian and B. Karasfi, "RDCGAN: Unsupervised representation learning with regularized deep convolutional generative adversarial networks," in *Proc. 9th Conf. Artif. Intell. Robot. 2nd Asia–Pacific Int. Symp.*, Dec. 2018, pp. 31–38, doi: 10.1109/AIAR.2018.8769811.

[27] H. Ham, T. J. Jun, and D. Kim, "Unbalanced GANs: Pre-training the generator of generative adversarial network using variational autoencoder," 2020, *arXiv:2002.02112*. [Online]. Available: http://arxiv.org/abs/2002.02112

CHUN-MING HUANG received the B.S., M.S., and Ph.D. degrees from the Department of Electrical Engineering, National Cheng Kung University, Taiwan, in 2000, 2005, and 2009, respectively. From 2010 to 2018, he was with Chung-Shan Institute of Science and Technology, Lungtan, Taiwan, as an Assistant Scientist. Since 2018, he has been a Faculty Member with National Formosa University, Yunlin, Taiwan, where he is currently an Assistant Professor with the Department of Electronic Engineering. His major research interests include error control codes and optical communications.

EDDY WIJANTO received the B.S. degree from the Department of Electrical Engineering, Krida Wacana Christian University, Indonesia, in 2005, and the M.S. degree from the Department of Electrical Engineering, Pelita Harapan University, Indonesia, in 2009. He is currently pursuing the Ph.D. degree in electro-optical engineering with National Formosa University, Yunlin, Taiwan. He is currently an Assistant Professor with the Department of Electrical Engineering, Krida Wacana Christian University. His major research interests include wireless and optical communications.

HSU-CHIH CHENG received the B.S. degree from the Department of Electronics, National Taiwan University of Science and Technology, in 2000, and the M.S. and Ph.D. degrees in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, in 2002 and 2006, respectively. He is currently a Full Professor with the Department of Electro-Optical Engineering, National Formosa University, Yunlin, Taiwan. His major research interests include DWDM networking devices, optical system design, and optics fiber sensor.

EDDY WIJANTO-FILE 3

ORIGINALITY REPORT

12%	9%	11%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	linknovate.com Internet Source	1%
2	blockgeni.com Internet Source	1%
3	eprints.utp.edu.my Internet Source	1%
4	aoot.osa.org Internet Source	1%
5	Submitted to BITS, Pilani-Dubai Student Paper	1%
6	affine.medium.com Internet Source	<1%
7	cyberleninka.org Internet Source	<1%
8	Adamu Abubakar Abba, Zhili Chen, Dong An. "Analysis and Correction of Micro CT Images for 3D Reconstruction", 2022 IEEE 13th International Conference on Software	<1%

	Engineering and Service Science (ICSESS), 2022 Publication	
9	myukk.org Internet Source	<1 %
10	researchr.org Internet Source	<1 %
11	Joao Jorge, Mauricio Villarroel, Sitthichok Chaichulee, Gabrielle Green, Kenny McCormick, Lionel Tarassenko. "Assessment of Signal Processing Methods for Measuring the Respiratory Rate in the Neonatal Intensive Care Unit", IEEE Journal of Biomedical and Health Informatics, 2019 Publication	<1 %
12	www.simplilearn.com Internet Source	<1 %
13	media.heidelbergengineering.com Internet Source	<1 %
14	Yi Zhou, Kai Yu, Meng Wang, Yuhui Ma, Yuanyuan Peng, Zhongyue Chen, Weifang Zhu, Fei Shi, Xinjian Chen. "Speckle Noise Reduction for OCT Images based on Image Style Transfer and Conditional GAN", IEEE Journal of Biomedical and Health Informatics, 2021 Publication	<1 %

15	ris.utwente.nl Internet Source	<1%
16	Fatemeh Mostafavi, Mohammad Tahsildoost, Zahra Sadat Zomorodian, Seyed Shayan Shahrestani. "An interactive assessment framework for residential space layouts using pix2pix predictive model at the early-stage building design", Smart and Sustainable Built Environment, 2022 Publication	<1%
17	Kenji Enomoto, Ken Sakurada, Weimin Wang, Hiroshi Fukui, Masashi Matsuoka, Ryosuke Nakamura, Nobuo Kawaguchi. "Filmy Cloud Removal on Satellite Imagery with Multispectral Conditional Generative Adversarial Nets", 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 Publication	<1%
18	Wen Bao, Yi Shen, Tao Chen, Peng Li, Zhihua Ding. "High-speed high-precision and ultralong-range complex spectral domain dimensional metrology", Optics Express, 2015 Publication	<1%
19	Yu-Sheng Tsai, Shun-Hsi Wang, Chuan-Hung Chen, Chien-Lung Cheng, Teh-Chao Liao. "Using copper substrate to enhance the thermal conductivity of top-emission organic	<1%

light-emitting diodes for improving the luminance efficiency and lifetime", Applied Physics Letters, 2009
Publication

20	drtc.isibang.ac.in Internet Source	<1 %
21	Hajer Walhazi, Ahmed Maalej, Najoua Essoukri Ben Amara. "Mask2LFP: Mask-constrained Adversarial Latent Fingerprint Synthesis", 2020 International Conference on Cyberworlds (CW), 2020 Publication	<1 %
22	Saira Osama, Kashif Zafar, Muhammad Usman Sadiq. "Predicting Clinical Outcome in Acute Ischemic Stroke Using Parallel Multi-parametric Feature Embedded Siamese Network", Diagnostics, 2020 Publication	<1 %
23	encyclopedia.pub Internet Source	<1 %
24	research-management.mq.edu.au Internet Source	<1 %
25	Anjing Guo, Leyuan Fang, Min Qi, Shutao Li. "Unsupervised Denoising of Optical Coherence Tomography Images with Nonlocal-Generative Adversarial Network",	<1 %

IEEE Transactions on Instrumentation and Measurement, 2020
Publication

26	Park, Seong Jun, Kwan Seob Park, Young Ho Kim, and Byeong Ha Lee. "Simultaneous Measurements of Refractive Index and Thickness by Spectral-Domain Low Coherence Interferometry Having Dual Sample Probes", IEEE Photonics Technology Letters, 2011. Publication	<1 %
27	Submitted to University of St Mark and St John Student Paper	<1 %
28	Yifeng Tang, Wanrong Gao. "Effects of orientation deviation of a beam splitter and a reference mirror on the stability of a two-interferometer-based handheld FFOCT imaging probe", Applied Optics, 2021 Publication	<1 %
29	inct.furg.br Internet Source	<1 %
30	tvst.arvojournals.org Internet Source	<1 %
31	www.repository.cam.ac.uk Internet Source	<1 %
32	Jiewen Lin, Shuncong Zhong, Qiukun Zhang, Weiqiang Chen. "Five-Frame Variable Phase-	<1 %

Shifting Method for Full-Range Spectral-Domain Optical Coherence Tomography", Applied Sciences, 2018
Publication

33	Lin An. "Full range complex spectral domain optical coherence tomography for volumetric imaging at 47000 A-scans per second", Journal of Optics, 08/01/2010 Publication	<1 %
34	Mufeng Geng, Xiangxi Meng, Lei Zhu, Zhe Jiang, Mengdi Gao, Zhiyu Huang, Bin Qiu, Yicheng Hu, Yibao Zhang, Qiushi Ren, Yanye Lu. "Triplet Cross-Fusion Learning for Unpaired Image Denoising in Optical Coherence Tomography", IEEE Transactions on Medical Imaging, 2022 Publication	<1 %
35	personales.upv.es Internet Source	<1 %

Exclude quotes	On	Exclude matches	< 15 words
Exclude bibliography	On		