RESEARCH ARTICLE

BRCycle-GAN: A Near-Infrared Fluorescence Image Processing Network Based on a Small Training Set

YURAN WANG, **LUGUI WANG**, AND **YE TIAN**
Biomaterials Research Center, School of Biomedical Engineering, Guangdong Provincial Key Laboratory of Construction and Detection in Tissue Engineering, Southern Medical University, Guangzhou 510515, China

Corresponding author: Ye Tian (tymty@smu.edu.cn)

ABSTRACT As a noninvasive, nonradiative and high-speed imaging modality, fluorescence imaging in the second near-infrared window (NIR-II, 1,000-1,700 nm) has demonstrated great potential for biomedical research and clinical study. The NIR-II window can be further divided into two spectral regions: NIR-IIa (1,000-1,300 nm) and NIR-IIb (1,500-1,700 nm). Compared to NIR-IIa, imaging in NIR-IIb region affords high-resolution imaging at subcentimeter tissue depths due to suppressed photon scattering and diminished tissue autofluorescence at long wavelengths, but relies on probes with high toxicity. To address the problem, researchers employ deep learning networks to attain NIR-IIb images from NIR-IIa images. However, current methods require numerous paired or unpaired images (more than 2800 images) as training sets, which can hardly acquire. In this work, an innovative convolutional neural network (BRCycle-GAN) is trained based on a small training set (merely 63 images) to transform NIR-IIa images into images with NIR-IIb imaging qualities. The NIR-IIb images generated by BRCycle-GAN outperform previous network models in terms of peak signal-to-noise ratio, cosine similarity and other image evaluation indices.

INDEX TERMS Deep learning, biomedical image processing, optical imaging.

I. INTRODUCTION

With enhanced penetration depth and improved signal-to-noise ratio (SNR), fluorescence imaging in the second near-infrared window (NIR-II, 1,000-1,700 nm) has emerged as a novel imaging modality for noninvasive, real-time imaging of internal structure and function in vivo [1], [2], [3], [4], [5], [6]. Benefiting from low tissue scattering and autofluorescence, imaging in the NIR-IIb window (1,500-1,700 nm) has been reported to achieve high spatial resolution deep inside living mammals when compared to NIR-IIa window (1,000-1,300 nm) and thus has attracted much attention in the field of biomedical research [7], [8], [9], [10], [11], [12]. Fluorescent probes that emit in the NIR-IIb

region are prerequisite for NIR-IIb imaging and the reported NIR-IIb probes are mainly based on inorganic quantum dots (QDs) [4], [13], [14] and rare earth nanoparticles (RENPs) [15], [16], [17]. Although NIR-IIb imaging of animal models can be successfully realized with these probes, their toxicity cannot be neglected in further clinical study. Therefore, many researchers turned to imaging processing that generates images with resolution and SNR approximating NIR-IIb images from NIR-IIa images [18], [19].

Conversion of NIR-IIa images to NIR-IIb images requires image style migration without any loss of original information. Recent years have witnessed rapid development of image processing in medical imaging, including transferring cross-modal images [20], [21], [22], denoising low-dose images [23], [24], [25] and reconstructing super-resolved microscopic images [26], [27]. Goodfellow et al. [28] firstly

---

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

proposed generative adversarial network (GAN) in 2014, which was a milestone in image processing and subsequently derived various GAN-based style migration models, such as pix2pix [29], cycle-GAN [19] and starGAN [30]. GAN is the most promising network model for unsupervised learning on complex distribution recently, but often suffers from convergence difficulty and generator collapse. Zhu et al [19] proposed cycle-GAN in 2017, which was a famous algorithm model for image-to-image translation with unpaired training sets [31]. Pix2pix employs U-net [32] model and achieves highly retaining of detailed information, while relies on strictly paired images for training.

Compared to traditional algorithms [18], deep learning demonstrates excellent performance in feature learning and mapping when processing fluorescence images. Ma [33] and coworkers employed cycle-GAN to convert images captured in NIR-I (800-1000 nm) and NIR-IIa windows into images with similar resolution and SNR to NIR-IIb images [19]. However, this network requires large amount of NIR-IIa and NIR-IIb images (more than 2800 images) for training, which are unavailable for average users. To address the problem, in this study, we proposed a modified cycle-GAN network, BRCycle-GAN, that requires a training set of merely 63 images and affords image quality higher than that obtained using previous reported methods. In BRCycle-GAN, a bottleneck residual module is introduced in the generators to increase the network depth and reduce loss of original information. The modification further improves image clarity as well as image evaluation parameters including peak SNR and cosine similarity. BRCycle-GAN lowers the threshold for users and thus broaden the application of NIR-II imaging in the field of biomedicines.

II. MATERIALS

A. DATABASE DESCRIPTION
The training data used in this study are 63 paired NIR-IIa and NIR-IIb 2D images with 512×640 pixels and a bit depth of 8. Each image is randomly sliced into 256×320 pixels before inputting into the network and a horizonal flip of 5º is employed to improve the diversity of training set. The test set is 20 paired images.

B. NIR-II FLUORESCENT PROBES
PbS/CdS core-shell quantum dots coated with oleyamine-branched polyacrylic acid are prepared according to previous reports [13] and then PEGlyated to obtain CSQD-OPA-PEG, which is employed as a fluorescent probe that emits in NIR-IIb region. Polyethylene glycol grafted polystyrene (PS-g-PEG) is synthesized according to previous reports [34] and a commercial dye, IR813 (I863886, Macklin), is encapsulated in a micelle formed with PS-g-PEG to obtain p-813 [35], which is employed as a fluorescent probe that emits in NIR-IIa region. Detailed structures, synthetic methods and optical properties can be found in references listed above.

C. MOUSE HANDLING
All animal experiments are approved by the ethics committee of Southern Medical University (SYXK(yue)2016-0167). C57 mice were purchased from Risemice Biotech. Hair of mice is carefully removed by Nair to avoid damage to skin before subjected to fluorescence imaging. For a typical imaging experiment, a mouse is anesthetized by inhalation of isoflurane (2.5%) under an air flow of 10 L min^{-1}. A continuous flow of isoflurane (2%) is applied to maintain the anesthesia status of the mouse. CSQD-OPA-PEG and p-813 are dispersed in PBS at an equivalent concentration of 1.5 mg mL^{-1} and the mixture (0.15 mL) is administered intravenously through tail vein. The mouse is allowed to fully recover before being returned to the cage.

D. WIDE-FIELD NIR-II FLUORESCENCE IMAGING
A wide-field setup for NIR-II fluorescence imaging is described in our previous reports [35]. In brief, a mouse injected with probes is mounted on the stage beneath a zoom lens (1-50502-IR, Navitar). The excitation light is provided by an 808 nm laser through a multi-mode optical fiber (F-MSC-C-1SMA, Newport) with an average excitation power density of 60 mW cm^{-2} at the imaging plane. A water-cooled InGaAs charge coupled device camera (NIRvana HS, Princeton Instruments) is used to record the image. Filters of 1000 nm longpass (FELH1000, Thorlabs) and 1300 nm shortpass (89-676, Edmund Optics) are used for imaging in NIR-IIa window and a 1500 nm longpass filter (FELH1500, Thorlabs) is used for imaging in NIR-IIb window. Images are saved in the form of 32-bit grayscale with a size of 512×640 pixels and transformed into 8-bit ones using ImageJ.

III. METHODS

A. NEURAL NETWORK ARCHITECURE
This study proposes a network model, BRCycle-GAN, consisting of two generators (G_{a2b}, G_{b2a}) and two discriminators (D_{a2b}, D_{b2a}). The generators realize mapping between NIR-IIa and NIR-IIb images and the discriminators determine whether the images are real or synthesized by generators. When training the model, we input real NIR-IIa and NIR-IIb images are inputted into the network and the generators generates synthesized images, which are randomly inputted into the discriminators together with the real ones. The discriminators recognize whether the inputted image is real or not and feed the result back to the generators to improve the quality of generated images until the discriminator cannot recognize the synthesized images. This process is achieved by the adversarial loss function L_{adv}, similar to a two-player game, where the generators minimize objective function and the discriminators maximize it until equilibrium is reached [19]. This is a complete cyclic process of mapping between NIR-IIa and NIR-IIb images, which is shown in Fig. 1A.
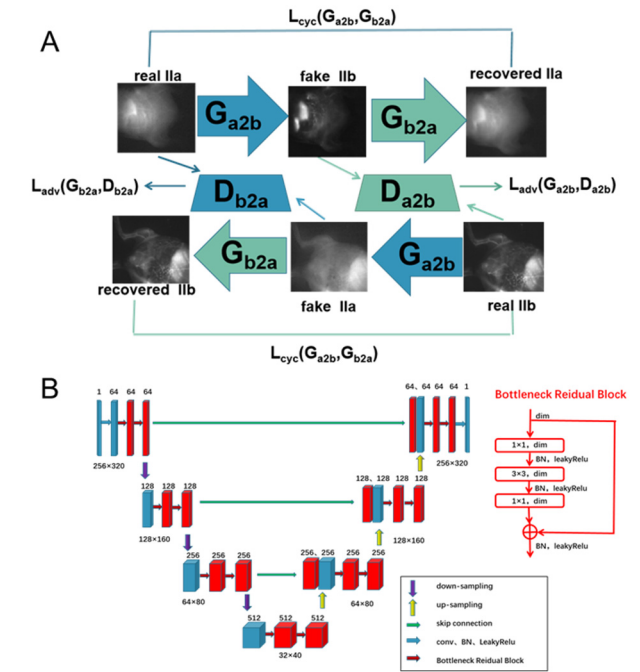
FIGURE 1. BRCycle-GAN-based NIR-IIa–to–NIR-IIb image transfer. (A) Training process of the BRCycle-GAN model. (B) Network architecture for generators.

B. LOSS FUNCTION

The goal is to learn the mapping function between NIR-IIa domain (denoted as A) and NIR-IIb domain (denoted as B) based on NIR-IIa images (denoted as $\{x_i\}_{i=1}^N$) and NIR-IIb images (denoted as $\{y_i\}_{i=1}^N$) in the training set. Generator G_{a2b} implements the mapping A→B and generator G_{b2a} implements the mapping B→A. Adversarial discriminator D_{a2b} aims at distinguishing real images $\{y\}$ from synthesized ones $\{G_{a2b}(x)\}$ and D_{b2a} aims at distinguishing real images $\{x\}$ from synthesized ones $\{G_{b2a}(y)\}$. Adversarial loss function L_{adv} is used to match generated image with data in the target domain and cycle consistent loss function L_{cyc} is used to ensure that the learned mapping is similar to original image.

The loss function is defined as follows:

$$L_{adv}(G_{a2b}, D_{b2a}) = E_{x\sim A}||D_{b2a}(G_{a2b}(x)) - 1||_2^2$$
$$+ E_{y\sim B}||D_{b2a}(y)||_2^2$$
$$L_{adv}(G_{b2a}, D_{a2b}) = E_{y\sim B}||D_{a2b}(G_{a2b}(y)) - 1||_2^2$$
$$+ E_{x\sim A}||D_{a2b}(x)||_2^2$$
$$L_{cyc}(G_{a2b}, G_{b2a}) = E_{x\sim A}||G_{b2a}(G_{a2b}(x)) - x||_1$$
$$+ E_{y\sim B}||G_{a2b}(G_{b2a}(y) - y||_1$$
$$L(G_{a2b}, G_{b2a}, D_{a2b}, D_{b2a}) = L_{adv}(G_{a2b}, D_{b2a})$$
$$+ L_{adv}(G_{b2a}, D_{a2b})$$
$$+ \Upsilon L_{cyc}(G_{a2b}, G_{b2a})$$

Adversarial loss L_{adv} is a mean squared error loss and cycle consistent loss L_{cyc} is a L1 loss. Total loss is a weighed sum of adversarial loss and cycle consistent loss. Hyperparameter $\Upsilon = 10$ is set for all experiments. Generators are trained to minimize the loss, while discriminators are trained to maximize the loss.

C. U-NET-LIKE GENERATOR

Compared to traditional cycle-GAN network using U-net [32] in generators, bottleneck residual blocks are introduced in generators of BRCycle-GAN (Fig. 1B) to increase network depth, facilitating extraction of deep image features and attenuation of gradient vanishing. The generator consists of encoding layers and decoding layers, which are connected by skip connection. Encoding layers acquire shallow features of the image and decoding layers acquire deep features, allowing generated image containing deep features without loss of shallow features.

Generators in this work have 49 convolution layers, including 4 encoding and 4 decoding layers. We extract 64 feature maps from 256×320 pixels 2D grayscale images with channel number 1 using 3×3 convolution kernels as inputs (stride = 1, padding = 1), followed by batch normalization and a leaky ReLu with 0.2 negative. Batch normalization, proposed by Peng et al. [36], provides several advantages, such as fast training, better performance and low sensitivity to initialization. Pooling, widely applied in classification or recognition tasks [37] to reduce dimensions of feature maps, often decreases sensitivity to small differences, loses image features and thus is not employed in BRCycle-GAN. Between encoding layers, down-sampling of feature maps is conducted, using 3×3 convolution kernels with stride of 2, padding of 1 and channel numbers are thus doubled after each down-sampling. Architecture of the following bottleneck residual block is shown in the right side of Fig. 1B: it consists of two 1×1 convolution layers and one 3×3 convolution layer and batch normalization of the feature map and a leaky ReLu with 0.2 negative are added between layers. After 4 encoding layers, obtained image features are sent to decoding layers for deconvolution, which utilize 2×2 convolution kernels (stride = 2, padding = 2). Shallow features from encoding layers and feature maps obtained from up-sampling are both inputted into two bottleneck residual blocks in decoding layers. After 4 decoding layers, generator produces a 256×320 pixels synthesized image.

D. DISCRIMINATOR ARCHITECTURE

Size of discriminator input is the same as that of generator output (256×320 pixels). PatchGAN [29] is used as the discriminator, which contains 5 convolution layers, each followed by a batch normalization (except for the first layer) and a leaky ReLu activation with a slope of 0.2 for negative input (except for the last layer). Convolution kernel is 4×4 with stride of 2, padding of 1 and channel numbers are thus doubled after each layer. Receptive field of final output is 70×70 pixels.

E. TRAINING PARAMETER

System configurations: Intel Core i9-10900X CPU HP, 3.7 GHz NVIDIA GeForce GTX A6000 GPU along with 40 GB RAM and Windows 10 (64 bit) operating system. Pytorch is used as the framework to train the models. All weights of neural networks are initialized with a normal distribution with a mean of 0 and a standard deviation of 0.02, and all bias terms are initialized as 0. To train the neural networks, we apply an Adam optimize [38] with hyperparameters $\beta_1 = 0.5$ and $\beta_2 = 0.999$. Learning rate is 0.0002 for 100 epochs. Training period is optimized: less iterations cannot generate meaningful images and too many iterations will overfit and produce artifacts. For ResNet generator, learning rate is 0.0002 for the first 50 epochs and then decay linearly to 0 in the next 50 epochs. A batch size of 4 is used during training. In each iteration, a random NIR-IIa image and a random NIR-IIb image are selected from training set and the loss is computed. Generated images are stored in a generated image pool [19]. The optimizer then updates parameters of generators. The loss of discriminators is calculated using the real image and a random image from generated image pool [19] and the optimizer updates parameters of discriminators [19]. After training, the generator G_{a2b} is used to transfer an input NIR-IIa image to an image that resemble a NIR-IIb image. Our code is publicly released on GitHub (https://gitee.com/wanglugui979899/wlg-brcycle-gan.git).

IV. EXPERIMENTS AND RESULTS

A. EVALUATION METHODS

In evaluating the quality of the images generated by our network, we use the ground-truth captured NIR-IIb window fluorescence images as the base truth values for calculating the corresponding evaluation metrics. In this study, mean square error (MSE), mean absolute error (MAE), peak signal-to-noise ratio (PSNR), and cosine similarity (COSINs) are used as evaluation metrics, where smaller MSE and MAE indicate better quality of the generated images, and larger PSNR and COSINs indicate better quality of the generated images. MAE denotes the average of the absolute errors of pixel values of the generated images and the real NIR-IIb images. The MAE is calculated as follows:

$$\text{MAE} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |x(i,j) - y(i,j)|$$

MSE denotes the average of the sum of squares of the difference between the pixel values of the generated image and the real NIR-IIb image, and the MSE is computed as follows:

$$\text{MSE} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x(i,j) - y(i,j)]^2$$

PSNR is an objective measure of the quality of the generated image and the de-noising ability of the real NIR-IIb image. PSNR is usually expressed in logarithmic decibel units and the PSNR is computed as follows:

$$\text{PSNR} = 10 \times \log_{10}\left(\frac{255^2}{\text{MSE}}\right)$$

Here, x(i, j) denotes the value of pixel (i, j) in the generating images y(i, j) represents the value of pixel (i, j) in the ground truth image, and m × n is the total number of pixels.

COSINs measures the similarity between the generated image and the real NIR-IIb image by measuring the cosine value of the inner product space of the two vectors of the generated image. The formula for COSINs is as follows:

$$COSINs = \sum_{i=0}^{n-1} (x_i \times y_i) \Bigg/ \left(\sqrt{\sum_{i=0}^{n-1} x_i^2} \times \sqrt{\sum_{i=0}^{n-1} y_i^2} \right)$$

B. QUALITATIVE EVALUATION OF IMAGE QUALITY

To compare generated images with real ones, NIR-IIa fluorescence probe, p-813, is injected intravenously into C57 mice together with NIR-IIb fluorescence probe, CSQD-OPA-PEG. In this case, wide field NIR-IIa and NIR-IIb images of the same view can be simultaneously collected. NIR-IIa images are inputted into the trained generator G_{a2b} and a new image is generated. The new image is not same as any image in training set, indicating no memory effect of the network [33]. When compared with original NIR-IIa one, the generated image exhibits high contrast and clarity with vasculature pattern preserved and sharpened (Fig. 2A). Signal-to-background ratio (SBR) of images are analyzed by plotting the cross-sectional intensity profiles of the same areas in the NIR-IIa, NIR-IIb and generated images. The generated image is very similar to real NIR-IIb image with highly overlapped fluorescence intensity curves (Fig. 2B). The generated images have the same high SBR as the NIR-IIb images, and both are higher than the NIR-IIa images (Fig. 2C). These results indicate that BRCycle-GAN can enhance contrast of NIR-IIa image precisely without introducing artifacts.

We further compare image processing of our BRCycle-GAN with traditional cycle-GAN [19] and pix2pix [29] models (both using U-net as generator). As shown in Fig. 3A, BRCycle-GAN produces image with the highest similarity to real NIR-IIb image. Image generated by cycle-GAN is grainy and noisy, while pix2pix generated blurry image with vessel detail loss. In the second row of Fig. 3A, which is a zoomed-in view of the ROI region, we can see that BRCycle-GAN generates an image with the sharpest details and the most similarity to the real NIR-IIb image, and pix2pix generates an image with blurred details and the strongest noise.

As reported previously, CSQD-OPA and IR813-OPA can effectively accumulate in bone marrow and thus can be used as fluorescence probes for bone imaging in NIR-IIb and NIR-IIa windows, respectively. NIR-IIa bone images are also processed with BRCycle-GAN. Image generated with BRCycle-GAN is significantly better than those generated by cycle-GAN and pix2pix (Fig. 3B). Fluorescence intensity
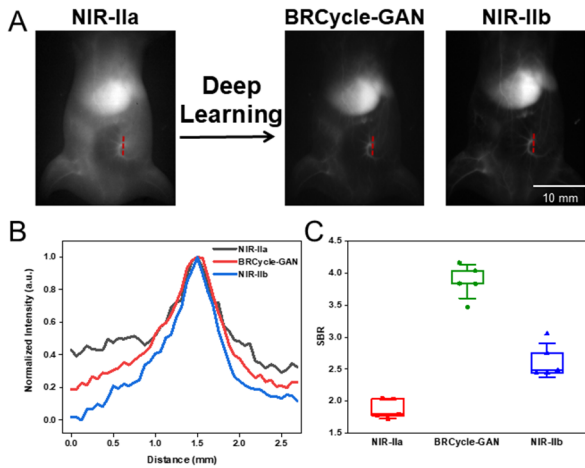
FIGURE 2. Wide-field fluorescence imaging with BRCyle-GAN. (A) In vivo fluorescence imaging of a C57 mouse injected with p-813 and CSQD-OPA-PEG. The NIR-IIa image was processed by the generator G_{a2b} to obtain the contrast-enhanced image. (Scale bar, 10mm.) (B) Cross-sectional intensity profiles of the same vessel (labeled in A) in the NIR-IIa, NIR-IIb, and generated images. (C) SBR of 5 NIR-IIa, NIR-IIb and generated images.
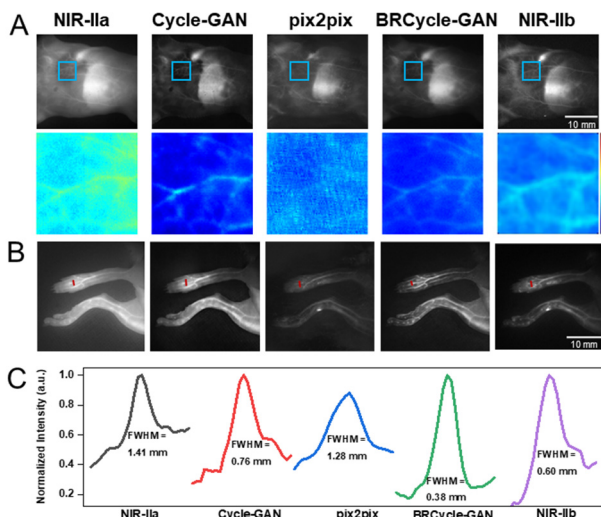

FIGURE 3. Comparison of images generated by different models. In (A), the first row displays wide-field fluorescence images of C57 mice generated by different models, the second row shows a magnified view of the vascular ROI. From left to right, each column presents NIR-IIa, Cycle-GAN, pix2pix, BRCycle-GAN, and NIR-IIb. (Scale bar, 10 mm.) (B) Illustrates a comparison of bone imaging comparisons. (C) Normalized fluorescence intensity of the red dotted line shown in B. Fluorescence intensity was normalized by the maximum intensity on the line.

curves of red dotted line in Fig. 3B are shown in Fig. 3C. Image generated by BRCycle-GAN is similar to real NIR-IIb image and has the smallest half-peak width. The half-peak width of the original NIR-IIa image is 1.41mm, while the half-peak width of the image generated by our model is only 0.38mm, which is reduced by 73.05%, while the half-peak width of the image generated by Cycle-GAN is 0.76mm, and the half-peak width of the image generated by pix2pix is 1.28mm, which are higher than that of the image generated

TABLE 1. Test results of Cycle-GAN, pix2pix and BCRCycle-GAN models.

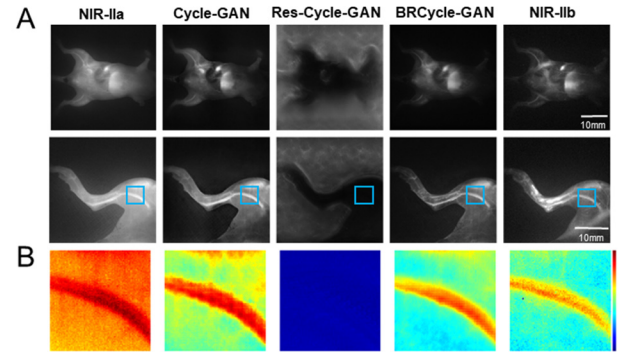| Methods	NIR-IIa	Cycle-GAN	pix2pix	BRCycle-GAN
PSNR(dB)	15.9440 ± 2.9151	19.7981 ± 4.6182	15.7513 ± 5.8382	**20.0481 ± 4.0043**
COSIN$_S$	0.6927 ± 0.2013	0.6990 ± 0.0982	0.6639 ± 0.1656	**0.7475 ± 0.1294**
MAE	0.1353 ± 0.0437	0.0895 ± 0.0328	0.0471 ± 0.0200	**0.0848 ± 0.0374**
MSE	0.0296 ± 0.0169	0.0146 ± 0.0075	0.0310 ± 0.0300	**0.0134 ± 0.0079**

FIGURE 4. Ablation experiment. In (A), the first and second row display wide-field fluorescence images of C57 mice generated by different models, from left to right, each column presents NIR-IIa, Cycle-GAN, Res-Cycle-GAN, BRCycle-GAN, and NIR-IIb. (Scale bar, 10 mm.) (B) shows a magnified view of the vascular ROI (labled in A).

by our model. These results indicate superiority of BRCycle-GAN in processing NIR-II bone images.

C. QUANTITATIVE EVALUATION

To compare image quality quantitatively, peak SNRs (PSNRs), cosine similarities (COSINs), MSEs and MAEs of 20 paired images in test set are calculated and shown in Table 1. Using the real NIR-IIb images as the reference, the PSNRs of the original NIR-IIa images, Cycle-GAN, pix2pix and BRCycle-GAN model generated images were 15.94±2.92 dB, 19.80±4.62 dB, 15.75±5.84 dB and 20.05±4.00 dB respectively. PSNRs of BRCycle-GAN generated images are higher than those generated by Cycle-GAN and pix2pix. The PSNRs of images generated by our model are 25.78% higher than those of original images and 27.30% higher than those generated by pix2pix. COSINs of our model are also best among these networks, from 0.6690±0.0982 and 0.6639±0.1656 for Cycle-GAN and pix2pix models to 0.7475±0.1294 for our model generated images. The COSINs is 6.94% higher than the existing optimal model Cycle-GAN.

We also further calculated MAE and MSE, and the images generated by BRCycle-GAN have smaller MAE values than the Cycle-GAN and original NIR-IIa images. In addition, the MSE value of the image generated by our model is the smallest among all models, which is 54.73% smaller than the original NIR-IIa image and 8.22% smaller than the best existing model. In summary, these results demonstrate excellent performance of BRCycle-GAN in processing wide-field NIR-II vascular images.

TABLE 2. Test results of Cycle-GAN, Res-Cycle-GAN and BRCycle-GAN models.

Methods	NIR-IIa	Cycle-GAN	Res-Cycle-GAN	BRCycle-GAN
PSNR(dB)	15.9440 ± 2.9151	19.7981 ± 4.6182	13.9906 ± 4.0448	**20.0481 ± 4.0043**
COSIN$_s$	0.6927 ± 0.2013	0.6990 ± 0.0982	0.6732 ± 0.2636	**0.7475 ± 0.1294**
MAE	0.1353 ± 0.0437	0.0895 ± 0.0328	0.1660 ± 0.0610	**0.0848 ± 0.0374**
MSE	0.0296 ± 0.0169	0.0146 ± 0.0075	0.0514 ± 0.0370	**0.0134 ± 0.0079**

D. ABLATION EXPERIMENT

In order to verify the effect of bottleneck residual blocks in our model, we perform ablation experiments. For the same training set, three Cycle-GAN models are trained, including traditional Cycle-GAN (U-net as generator), Res-Cycle-GAN (ResNet as generator) and BRCycle-GAN (combining U-net and ResNet). As shown in Fig. 4A, Res-Cycle-GAN only learns outline and introduces numerous artifacts and noise. Cycle-GAN learns some image features, but the generated image is granular with high background and noise. In contrast, BRCycle-GAN generates realistic images with high SNR. In Fig. 4B, zooring the ROI we can clearly see that the images generated by our model are most similar to the real NIR-IIb images.As shown in Table 2, image generated by BRCycle-GAN exhibits the highest PSNR, 1.26% higher than tranditional Cycle-GAN and 43.30% higher than Res-Cycle-GAN. COSINs of BRCycle-GAN is also 6.94% higher than Cycle-GAN and 12.60% higher than Res-Cycle-GAN. The MAE and MSE values of the images generated by BRCycle-GAN are also the smallest among all models, which are 5.25% and 70.00% lower than Cycle-GAN and Res-Cycle-GAN for MAE values, respectively. For MSE, our models were 8.22% and 73.93% lower than Cycle-GAN and Res-Cycle-GAN respectively. Above results show that introduction of bottleneck residual blocks in U-net generator has a significant effect on improving PSNRs, COSINs of generated images, producing image with high quality.

V. CONCLUSION

In this paper, we propose an innovative network model, BRCycle-GAN, for NIR-II image processing based on a small training set. Bottleneck residual blocks are introduced in U-net generator of Cycle-GAN and network adjustments, such as discarding pooling layers and utilizing batch normalization instead of layer normalization, improve learning capability of the network, achieving high-quality conversion of NIR-IIa images to quasi-NIR-IIb images. Compared with existing NIR image conversion models, BRCycle-GAN demonstrates outstanding performance in processing NIR-IIa images. Ablation studies further confirm the outstanding performance is attributed to bottleneck residual blocks. Our innovative network model extends application of deep-learning based image processing in NIR-II imaging and broadens horizons for the design of novel algorithms for image processing.

REFERENCES

[1] M. Smith, C. Mancini, and S. Nie, "Bioimaging: Second window for in vivo imaging," *Nature Nanotechnol.*, vol. 4, no. 11, pp. 710–711, Nov. 2009.

[2] G. Hong, J. C. Lee, J. T. Robinson, U. Raaz, L. Xie, N. F. Huang, J. P. Cooke, and H. Dai, "Multifunctional in vivo vascular imaging using near-infrared II fluorescence," *Nature Med.*, vol. 18, no. 12, pp. 1841–1846, Dec. 2012.

[3] J. T. Robinson, G. Hong, Y. Liang, B. Zhang, O. K. Yaghi, and H. Dai, "In vivo fluorescence imaging in the second near-infrared window with long circulating carbon nanotubes capable of ultrahigh tumor uptake," *J. Amer. Chem. Soc.*, vol. 134, no. 25, pp. 10664–10669, Jun. 2012.

[4] G. Hong, J. Robinson, Y. Zhang, S. Diao, A. Antaris, Q. Wang, and H. Dai, "In vivo fluorescence imaging with Ag$_2$S quantum dots in the second near-infrared region," *Angew. Chem. Int. Ed.*, vol. 51, no. 39, pp. 9818–9821, Sep. 2012.

[5] C. Li, Y. Zhang, M. Wang, Y. Zhang, G. Chen, L. Li, D. Wu, and Q. Wang, "In vivo real-time visualization of tissue blood flow and angiogenesis using Ag$_2$S quantum dots in the NIR-II window," *Biomaterials*, vol. 35, no. 1, pp. 393–400, Jan. 2014.

[6] F. Hu, C. Li, Y. Zhang, M. Wang, D. Wu, and Q. Wang, "Real-time in vivo visualization of tumor therapy by a near-infrared-II Ag$_2$S quantum dot-based theranostic nanoplatform," *Nano Res.*, vol. 8, no. 5, pp. 1637–1647, Apr. 2015.

[7] S. Diao, J. L. Blackburn, G. Hong, A. L. Antaris, J. Chang, J. Z. Wu, B. Zhang, K. Cheng, C. J. Kuo, and H. Dai, "Fluorescence imaging in vivo at wavelengths beyond 1500 nm," *Angew. Chem.*, vol. 127, no. 49, pp. 14971–14975, Dec. 2015.

[8] Z. Luo, D. Hu, D. Gao, Z. Yi, H. Zheng, Z. Sheng, and X. Liu, "High-specificity in vivo tumor imaging using bioorthogonal NIR-IIb nanoparticles," *Adv. Mater.*, vol. 33, no. 49, Dec. 2021, Art. no. e210295.

[9] Y. Zhong, Z. Ma, F. Wang, X. Wang, Y. Yang, Y. Liu, X. Zhao, J. Li, H. Du, M. Zhang, Q. Cui, S. Zhu, Q. Sun, H. Wan, Y. Tian, Q. Liu, W. Wang, K. C. Garcia, and H. Dai, "In vivo molecular imaging for immunotherapy using ultra-bright near-infrared-IIb rare-Earth nanoparticles," *Nature Biotechnol.*, vol. 37, no. 11, pp. 1322–1331, Nov. 2019.

[10] N. Liao, L. Su, Y. Zheng, B. Zhao, M. Wu, D. Zhang, H. Yang, X. Liu, and J. Song, "In vivo tracking of cell viability for adoptive natural killer cell-based immunotherapy by ratiometric NIR-II fluorescence imaging," *Angew. Chem., Int. Ed.*, vol. 60, no. 38, pp. 20888–20896, Sep. 2021.

[11] M. Jiang, Z. Xue, and S. Zeng, "808 nm light triggered lanthanide nanoprobes with enhanced down-shifting emission beyond 1500 nm for imaging-guided resection surgery of tumor and vascular visualization," *Theranostics*, vol. 10, no. 15, pp. 6875–6885, 2020.

[12] Y. Li, S. Zeng, and J. Hao, "Non-invasive optical guided tumor metastasis/vessel imaging by using lanthanide nanoprobe with enhanced down-shifting emission beyond 1500 nm," *ACS Nano*, vol. 13, no. 1, pp. 248–259, Jan. 2019.

[13] M. Zhang, J. Yue, R. Cui, Z. Ma, H. Wan, F. Wang, S. Zhu, Y. Zhou, Y. Kuang, Y. Zhong, D.-W. Pang, and H. Dai, "Bright quantum dots emitting at ~1,600 nm in the NIR-IIb window for deep tissue fluorescence imaging," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 26, pp. 6590–6595, Jun. 2018.

[14] O. T. Bruns et al., "Next-generation in vivo optical imaging with short-wave infrared quantum dots," *Nature Biomed. Eng.*, vol. 1, no. 4, p. 0056, Apr. 2017.

[15] D. J. Naczynski, M. C. Tan, M. Zevon, B. Wall, J. Kohl, A. Kulesa, S. Chen, C. M. Roth, R. E. Riman, and P. V. Moghe, "Rare-Earth-doped biological composites as in vivo shortwave infrared reporters," *Nature Commun.*, vol. 4, no. 1, p. 2199, Jul. 2013.

[16] R. Wang, X. Li, L. Zhou, and F. Zhang, "Epitaxial seeded growth of rare-earth nanocrystals with efficient 800 nm near-infrared to 1525 nm short-wavelength infrared downconversion photoluminescence for in vivo bioimaging," *Angew. Chem. Int. Ed.*, vol. 53, no. 45, pp. 12086–12090, Sep. 2014, doi: 10.1002/ANIE.201407420.

[17] G. Hong, A. L. Antaris, and H. Dai, "Near-infrared fluorophores for biomedical imaging," *Nature Biomed. Eng.*, vol. 1, no. 1, p. 0010, Jan. 2017.

[18] D. Chen, W. Qi, Y. Liu, Y. Yang, T. Shi, Y. Wang, X. Fang, Y. Wang, L. Xi, and C. Wu, "Near-infrared II semiconducting polymer dots: Chain packing modulation and high-contrast vascular imaging in deep tissues," *ACS Nano*, vol. 17, no. 17, pp. 17082–17094, Aug.17, 2023.

[19] G. Wang, H. Shi, Y. Chen, and B. Wu, "Unsupervised image-to-image translation via long-short cycle-consistent adversarial networks," *Appl. Intell.*, vol. 53, no. 14, pp. 17243–17259, Jul. 2023.

[20] D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, and D. Shen, "Medical image synthesis with deep convolutional adversarial networks," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 9, p. 2706, Sep. 2020.

[21] H. Choi and D. S. Lee, "Generation of structural MR images from amyloid PET: Application to MR-less quantification," *J. Nucl. Med.*, vol. 59, no. 7, pp. 1111–1117, Jul. 2018.

[22] C.-B. Jin, H. Kim, M. Liu, W. Jung, S. Joo, E. Park, Y. Ahn, I. Han, J. Lee, and X. Cui, "Deep CT to MR synthesis using paired and unpaired data," *Sensors*, vol. 19, no. 10, p. 2361, May22, 2019.

[23] Y. Wang, B. Yu, L. Wang, C. Zu, D. S. Lalush, W. Lin, X. Wu, J. Zhou, D. Shen, and L. Zhou, "3D conditional generative adversarial networks for high-quality PET image estimation at low dose," *NeuroImage*, vol. 174, pp. 550–562, Jul. 2018.

[24] M. Gholizadeh-Ansari, J. Alirezaie, and P. Babyn, "Low-dose CT denoising with dilated residual network," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Honolulu, HI, USA, Jul. 2018, pp. 5117–5120.

[25] M. Ran, J. Hu, Y. Chen, H. Chen, H. Sun, J. Zhou, and Y. Zhang, "Denoising of 3D magnetic resonance images using a residual encoder–decoder Wasserstein generative adversarial network," *Med. Image Anal.*, vol. 55, pp. 165–180, Jul. 2019.

[26] H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydın, L. A. Bentolila, C. Kural, and A. Ozcan, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nature Methods*, vol. 16, no. 1, pp. 103–110, Jan. 2019.

[27] L. Jin, B. Liu, F. Zhao, S. Hahn, B. Dong, R. Song, T. C. Elston, Y. Xu, and K. M. Hahn, "Deep learning enables structured illumination microscopy with low light levels and enhanced speed," *Nature Commun.*, vol. 11, no. 1, p. 1934, Apr. 2020.

[28] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020.

[29] N. U. Islam, S. Lee, and J. Park, "Accurate and consistent image-to-image conditional adversarial network," *Electronics*, vol. 9, no. 3, p. 395, Feb. 2020.

[30] H. Tang, H. Liu, and N. Sebe, "Unified generative adversarial networks for controllable image-to-image translation," *IEEE Trans. Image Process.*, vol. 29, pp. 8916–8929, 2020.

[31] B. Wang, G. Ma, and M. Zhu, "Fast momentum contrast learning for unsupervised person re-identification," *IEEE Signal Process. Lett.*, vol. 28, pp. 2073–2077, 2021.

[32] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-Net for medical image segmentation," *J. Med. Imag.*, vol. 6, no. 1, Jan. 2019, Art. no. 014006.

[33] Z. Ma, F. Wang, W. Wang, Y. Zhong, and H. Dai, "Deep learning for in vivo near-infrared imaging," *Proc. Nat. Acad. Sci. USA*, vol. 118, no. 1, Jan. 2021, Art. no. e202144611.

[34] H. Wan, J. Yue, S. Zhu, T. Uno, X. Zhang, Q. Yang, K. Yu, G. Hong, J. Wang, L. Li, Z. Ma, H. Gao, Y. Zhong, J. Su, A. L. Antaris, Y. Xia, J. Luo, Y. Liang, and H. Dai, "A bright organic NIR-II nanofluorophore for three-dimensional imaging into biological tissues," *Nature Commun.*, vol. 9, no. 1, p. 1171, Mar. 2018.

[35] P. Zhang, R. Guo, H. Zhang, W. Yang, and Y. Tian, "Fluoropolymer coated DNA nanoclews for volumetric visualization of oligonucleotides delivery and near infrared light activated anti-angiogenic oncotherapy," *Adv. Sci.*, vol. 10, no. 32, Nov. 2023, Art. no. 2304633.

[36] H. Peng, Y. Yu, and S. Yu, "Re-thinking the effectiveness of batch normalization and beyond," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 1, pp. 465–478, Jan. 2024.

[37] L. Xiang, Y. Qiao, D. Nie, L. An, W. Lin, Q. Wang, and D. Shen, "Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI," *Neurocomputing*, vol. 267, pp. 406–416, Dec. 2017.

[38] V. Kungurtsev and F. Rinaldi, "A zeroth order method for stochastic weakly convex optimization," *Comput. Optim. Appl.*, vol. 80, no. 3, pp. 731–753, Dec. 2021.
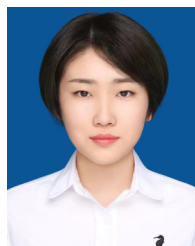
YURAN WANG received the bachelor's degree in biomedical engineering from Southern Medical University, China, in 2021, where she is currently pursuing the master's degree with the School of Biomedical Engineering.

Her research interests include near-infrared fluorescence imaging, deep learning, and cardiac electrophysiological detection.

LUGUI WANG was born in Chengmai County, Hainan, China, in November 1997. He received the bachelor's degree from the School of Biomedical Engineering, Southern Medical University, China, in 2020, where he is currently pursuing the degree with the School of Biomedical Engineering.

His research interests include image processing and deep learning, and in particular, he is now working on deep learning-based NIR fluorescence image processing techniques.

YE TIAN received the Ph.D. degree in polymer chemistry and physics from Fudan University, China, in 2017.

From 2017 to 2019, he was a Post-Doctoral Researcher at the Department of Chemistry, Stanford University, USA. In 2020, he joined Southern Medical University, China. He is currently a Professor with the School of Biomedical Engineering, Southern Medical University. His primary research interests include optical imaging, biomaterials, and neuroscience. Especially, he is now working on extending the toolbox of neuroscience study with novel optical imaging modality.

• • •