Enhancing Portable OCT Image Quality via GANs for AI-Based Eye Disease Detection

Kaveri A. Thakoor[1(✉)], Ari Carter[1], Ge Song[2], Adam Wax[3], Omar Moussa[4], Royce W. S. Chen[4], Christine Hendon[1], and Paul Sajda[1]

[1] Columbia University, New York, NY 10027, USA
{k.thakoor,ac4647,cpf2115,ps629}@columbia.edu
[2] University of Rochester, Rochester, NY 14627, USA
gsong4@u.rochester.edu
[3] Duke University, Durham, NC 27708, USA
a.wax@duke.edu
[4] Columbia University Medical Center, New York, NY 10032, USA
{om2333,rc2631}@cumc.columbia.edu

Abstract. Optical coherence tomography (OCT) is widely used for detection of ophthalmic diseases, such as glaucoma, age-related macular degeneration (AMD), and diabetic retinopathy. Using a low-coherence-length light source, OCT is able to achieve high axial resolution in biological samples; this depth information is used by ophthalmologists to assess retinal structures and characterize disease states. However, OCT systems are often bulky and expensive, costing tens of thousands of dollars and weighing on the order of 50 pounds or more. Such constraints make it difficult for OCT to be accessible in low-resource settings. In the U.S. alone, only 15.3% of diabetic patients meet the recommendation of obtaining annual eye exams; the situation is even worse for minority/under-served populations. In this study, we focus on data acquired with a low-cost, portable OCT (p-OCT) device, characterized by lower resolution, scanning rate, and imaging depth than a commercial OCT system. We use generative adversarial networks (GANs) to enhance the quality of this p-OCT data and then assess the impact of this enhancement on downstream performance of artificial intelligence (AI) algorithms for AMD detection. Using GANs trained on simulated p-OCT data generated from paired commercial OCT data degraded with the point spread function (PSF) of the p-OCT device, we observe improved AI performance on p-OCT data after single-image super-resolution. We also achieve denoising after image-to-image translation. By exhibiting proof-of-principle AI-based AMD detection even on low-quality p-OCT data, this study stimulates future work toward low-cost, portable imaging+AI systems for eye disease detection.

Supported by National Science Foundation Grant DGE 1644869.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-18523-6_15.

S. Albarqouni et al. (Eds.): DeCaF 2022/FAIR 2022, LNCS 13573, pp. 155–167, 2022.
https://doi.org/10.1007/978-3-031-18523-6_15

Keywords: Portable optical coherence tomography · Generative adversarial networks · Eye disease

1 Introduction

1.1 Background and Motivation

Much of the Artificial Intelligence (AI) being developed to detect ophthalmic diseases is trained using data and ground-truth that are collected in leading clinics with state-of-the-art equipment and expert ophthalmologists [1,2]. However, perhaps one of the most compelling uses for AI is for under-served areas which must utilize low-cost portable systems and in which high-quality ground truth may not be available given lack of experienced readers [3]. In this paper, we investigate whether we can use generative adversarial networks (GANs) to map the lower quality of data acquired using a portable-OCT (p-OCT) system (about tenfold cheaper than a commercial system) to the higher quality of commercial OCT data on which an AI model was trained to detect AMD. We hypothesize that this mapping will enable the p-OCT data to be "rescaled" to match the resolution and noise characteristics of the high-quality data used to train an AI model, enabling better downstream AI-based AMD classification performance. To test this, we utilize data collected from a p-OCT device developed by Kim *et al.* [4], who have successfully designed, implemented, and characterized a low-cost, portable OCT system tailored for retinal imaging use in clinical/laboratory studies. This p-OCT device was used for imaging human patients and achieved a contrast-to-noise ratio 5.6% less than that achieved by a commercial Heidelberg Spectralis OCT system [5].

Furthermore, since data shortage is a challenge for the development of AI-based tools, future potential applications of p-OCT data/devices include (1) efficient data collection and augmentation for AI training due to the p-OCT's portable form factor, and (2) GAN-based simulation/synthesis of medical data (as shown to be possible in previous work with commercial OCT and fundus images [6,7]) for enhancing data privacy. The above use cases of p-OCT data and GANs are predicated on the assumption of having robust AI algorithms that can achieve high-accuracy eye disease detection even from p-OCT data; we seek to confirm this assumption through our work by using GANs to improve downstream AI performance on p-OCT data. By showing proof-of-principle AI-based AMD detection using low-quality p-OCT data, we aim to equip a broader, diverse population with access to potentially sight-saving imaging+AI technology.

1.2 Past Work

Existing noise removal approaches for natural images have been applied to medical images [8,9] to successfully generate noise-free images even when clean target images are not available for training. In contrast, in this work we specifically seek to map low-quality image statistics of test data to high-quality image

statistics to mimic the quality of data on which an existing AI-based AMD detection model has been trained. This requires capturing the resolution difference between low and high-quality data (achieving super-resolution) as well as capturing the spatially-correlated noise differences between low and high-quality data (past approaches [9] often assume spatially uncorrelated/'pixel-wise independent' noise). In order to achieve this, we turn to the class of generative adversarial networks (GANs), particularly conditional GANs, which learn mappings between source and target data. First, we use super-resolution GANs (such as ESRGAN [10]) to enhance the resolution of output images by learning both the scale and noise mapping between low and high-quality data; such GANs enable 'informed hallucination' of missing information to generate super-resolved versions of input low-resolution data. Second, we use image-to-image translation GANs [11] (such as MedGAN [12]) for the denoising of simulated p-OCT images, to attempt to transform their perceptual quality toward that of commercial OCT images.

1.3 Baseline Performance on p-OCT Data and Dataset Details

Through a collaboration with Duke University's Wax Lab, we obtained 221 b-scan images from 59 subjects that were acquired directly with a p-OCT system [5]. We used this data (Institutional Review Board exempt) to test the authors' previously-developed [13,14] deep learning algorithm's (DLA's) robustness to AMD detection using p-OCT input images. (This previous DLA was trained from scratch on 301 commercial OCT b-scan images [14]). We showed the p-OCT images to AMD experts to label for presence of CNV (choroidal neovascularization, characteristic of 'wet AMD') or no AMD; 39 images were excluded due to presence of pathologies other than AMD, and 14 'non-neovascular' ('dry AMD') eyes were excluded in order to directly assess the DLA's binary AMD classification performance with p-OCT data vs. with commercial OCT data. This resulted in 168 p-OCT images (42 classified as neovascular (NV) AMD ('wet AMD') and 126 classified as non-AMD). For training, we utilized a dataset (described in detail in [14]) of 1270 NV AMD vs. non-AMD commercial OCT images (520 NV AMD, 750 non-AMD) captured with a Carl Zeiss Cirrus HD-OCT 5000 device. Thus, the training set is much closer to balanced (40% NV AMD vs. 60% non-AMD) even though the test set is not balanced. Commercial OCT training data and p-OCT test data examples are shown in Fig. 1 (left). Although both image types are impacted by characteristic speckle noise from the OCT instrument, the amount of noise and the distribution of pixel intensities vary significantly between the two images.

To present the baseline performance of the high-performing DLA in a threshold-independent manner (and to control for imbalance in classes, especially relevant for the p-OCT test data), we plotted the AMD detection model's performance via Receiver Operating Characteristic (ROC) curves (Fig. 3, left panel). The Area Under the Curve (AUC) achieved by the model on baseline **commercial OCT** data is **0.8 (95% CI, 0.694–0.906)**, while the baseline AUC on **p-OCT** data prior to any image quality enhancement is **0.518 (95%**

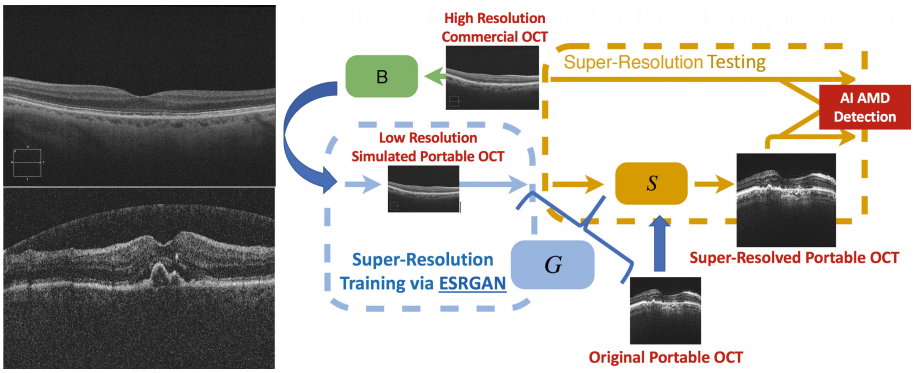

Fig. 1. Left: Original commercial training data (top); histogram-matched p-OCT test data (bottom). **Right:** Image processing pipeline to generate super-resolved images for downstream AI performance evaluation. Commercial OCT data is convolved with the PSF of the p-OCT device, histogram-matched with p-OCT images, and then downsampled by 4x (to form simulated p-OCT data through 'B') prior to being used for paired ESRGAN training. The resulting trained generator ('G') is used to super-resolve ('S') p-OCT test inputs prior to AI-based AMD detection (figure concept from [15]).

CI, 0.396–0.640), barely above chance. To improve the AI model's performance and make it generalizable for p-OCT data, we attempted one GAN-based super-resolution (SR) approach and one GAN-based image-to-image translation approach; these approaches are described in detail in the following sections.

2 Super-Resolving p-OCT Data with ESRGAN

We used the Enhanced Super Resolution Generative Adversarial Network (ESRGAN) [10], one of the latest state-of-the-art deep learning based super-resolution (SR) techniques, to enhance the resolution of the portable OCT test data prior to downstream classification via a high-performing AMD detection model to evaluate the impact of super-resolution on downstream binary AMD detection.

2.1 ESRGAN Background and Methods

ESRGAN goes beyond the Super-Resolution Generative Adversarial Network (SRGAN) [16] by creating the Enhanced SRGAN. Architectural highlights of ESRGAN that enable its enhanced functionality include the Residual-in-Residual Dense Block (RRDB) [10], which has higher capacity and thus is easier to train than the original SRGAN model, residual scaling [17], and use of a relativistic generator [18]. ESRGAN must be trained with paired low-resolution and high-resolution data (i.e. the same object captured with two imaging instruments). Since we do not have p-OCT and commercial OCT data of the same patients, we simulated paired ESRGAN training data by matching the histogram

of the commercial OCT data to that of the p-OCT data, convolving the commercial OCT data with the Point Spread Function (PSF) of the portable OCT, and downsampling the data by a factor of 4 (as dictated by the ESRGAN architecture [10]). This approach resembles past 'Blind Super-resolution' approaches [15] and is appropriate for this situation since we know the point spread function (i.e. degradation kernel) that generated the low-resolution p-OCT data. The pipeline used for generating the paired training data is shown in Fig. 1 (right). Examples of the degraded and ground truth commercial images used for paired training are shown in Fig. 2. We then super-resolved the p-OCT test data using the resulting ESRGAN generator trained (via BasicSR [19]) on the paired, 'simulated p-OCT'/commercial OCT dataset. After histogram-matching the p-OCT test data to the commercial OCT data and denoising via a 5×5 kernel Weiner filter, we inputted the p-OCT data to the trained ESRGAN generator for super-resolution. We then used the super-resolved outputs for downstream AI-based AMD detection.

2.2 ESRGAN Results and Discussion

Training ESRGAN on this paired dataset exhibits qualitatively improved results on validation data, as can be seen by the input and validation output images in Fig. 2. For test data (super-resolved p-OCT data), the improvement is visualized most clearly through the ROC curve in Fig. 3, left panel. AUC achieved by the high-performing AI model on baseline **commercial OCT** data is **0.8 (95% CI, 0.694–0.906)**, while baseline AUC on **portable OCT** data prior to application of super-resolution is **0.518 (95% CI, 0.396–0.640)**. In contrast, after ESRGAN super-resolution (without training on paired commercial OCT and simulated p-OCT data, i.e. just using an 'off-the-shelf' ESRGAN model pre-trained on natural images, called **After SR, No Train**), this AUC increases to **0.792 (95% CI, 0.684–0.900)**. After ESRGAN training on paired simulated p-OCT data and commercial OCT data (called **After SR, Train**), AI performance on super-resolved p-OCT data increases to an AUC of **0.897 (95% CI, 0.815–0.979)**. This increase in AUC beyond that achieved on the original commercial OCT data could be due to the super-resolution process in fact increasing the resolution of the p-OCT data beyond that of the original commercial OCT data (the ESRGAN architecture enables a $4\times$ resolution increase compared to input resolution; input resolution is 500×500 pixels, so output resolution is 2000×2000 pixels, while commercial OCT data resolution is only 700×1052 pixels).

It is interesting to note that although the AI performance on super-resolved p-OCT data improves dramatically compared to that on the commercial OCT data and that on the original p-OCT data, this improvement is harder to observe visually/perceptually (see Fig. 4). This is confirmed by computing the BRISQUE scores [20] for sets of these images. The BRISQUE score is a reference-less perceptual image quality metric; lower BRISQUE scores indicate higher perceptual quality. Figure 3 (right panel) shows that, while **commercial OCT** data and

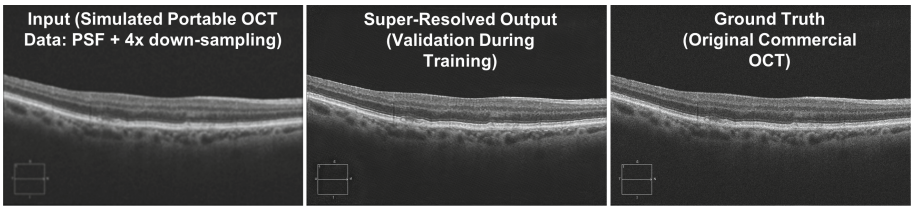

Fig. 2. Left: Simulated portable (degraded commercial) OCT data used for ESRGAN paired training, after histogram matching to p-OCT data, convolving with PSF of the p-OCT device, and 4x bicubic down-sampling; **Center:** SR validation output during training; **Right:** original commercial OCT ground truth.

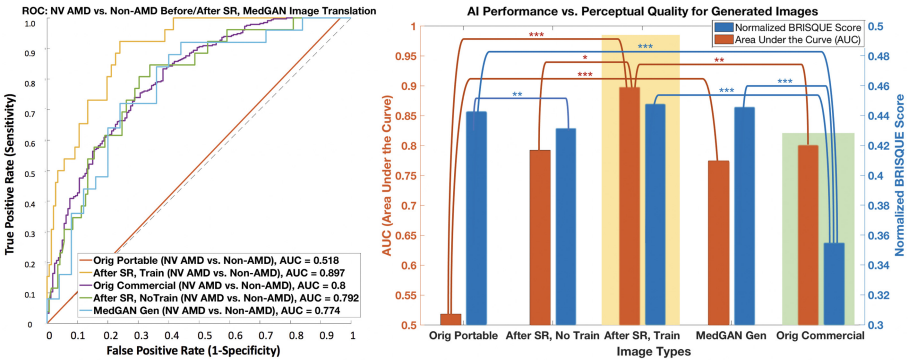

Fig. 3. Left Panel: Receiver Operator Characteristic (ROC) curve comparing performance of AI on **original p-OCT data** (red), on p-OCT data after super-resolution with paired training (**After SR, Train** in yellow), on **commercial OCT data** (violet), and on p-OCT data after super-resolution without training (**After SR, No-Train** in green). Interestingly, super-resolution enhances AI AUC beyond that on the high-quality commercial data, suggesting the value of super-resolution for facilitating AI-based eye disease detection even from p-OCT data. ROC curve in blue shows **MedGAN generated** images (described in Sect. 3) with AUC approaching that of target (commercial OCT) images. **Right Panel:** AUCs and normalized BRISQUE scores for 5 image types. Note inverse relationship between AUC scores and normalized BRISQUE scores for **After SR, Train** (highlighted with light yellow background) and **MedGAN Gen** images; they have significantly higher AUCs (AI performance) compared to **p-OCT images** but significantly poorer (higher) normalized BRISQUE scores (perceptual quality) compared to **commercial OCT** images (highlighted with light green background) [*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.0001$]. Note overall similarity in normalized BRISQUE scores across all image types. (Color figure online)

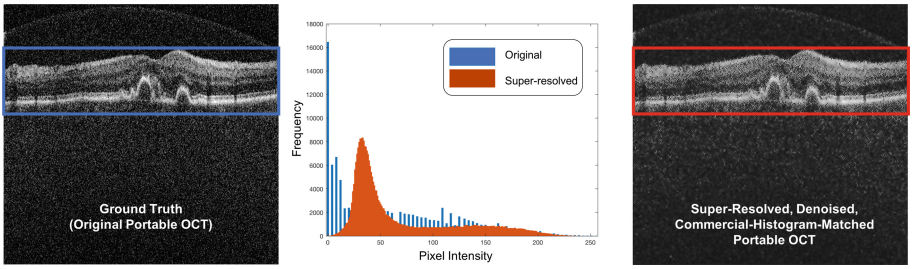

Fig. 4. Left: Original portable OCT image; **Right:** super-resolved, denoised, commercial-OCT-histogram-matched portable OCT image; **Center:** histogram of pixel intensity frequency quantitatively shows super-resolution (increase in presence of varied pixel intensities for red histogram, corresponding to red boxed region in super-resolved (SR) p-OCT image, compared to blue histogram, corresponding to blue boxed region in original p-OCT image). Note shifted peak of SR p-OCT histogram is due to learning histogram of commercial OCT data during paired ESRGAN training. (Color figure online)

super-resolved p-OCT data without training (**After SR, No Train**) have significantly different BRISQUE scores from that of the original **p-OCT** data, the trained, super-resolved (SR) p-OCT data (**After SR, Train**) data has a BRISQUE score that is *not significantly different* from that of the original **p-OCT** data. This aligns with the fact that clinicians who viewed the trained SR p-OCT data also qualitatively described that the trained SR p-OCT data did not provide any additional features beyond what the original p-OCT data provided for their diagnosis. This qualitative finding and the similar BRISQUE scores for the original p-OCT and trained SR p-OCT data, combined with the enhanced AI performance on the SR p-OCT data, are consistent with past work [21], which showed that deep neural networks (DNNs) are less impaired than their human counterparts at deciphering spatially correlated noise. Also, 'noise-trained' DNNs, like the trained ESRGAN here, more closely emulate human vision's robustness to noise than DNNs not trained with noise [21]. This further strengthens the potential value of an SR-enhanced, AI-embedded p-OCT system that could provide high-accuracy automated disease detection, *especially* when a human expert or commercial OCT system are not available, such as in resource-limited environments. The AI's boosted performance on the SR p-OCT data could be attributed to the SR process transforming the original p-OCT data into a space that is more similar to that on which the AMD-detection model has been trained (commercial OCT data). The effective 'transfer learning' during the paired training between simulated p-OCT data and true commercial OCT data also improved AI performance for the trained SR model compared to the non-trained one (when the SR model was only pre-trained on natural images).

3 Enhancing Source Domain Perceptual Image Quality with MedGAN

Given the dramatic improvement in AI performance with super-resolved p-OCT data without accompanying significant improvement in perceptual quality, we also implemented (in Python Tensorflow/Keras) an image-to-image translation [11] GAN to enhance p-OCT perceptual image quality via image denoising.

3.1 MedGAN Background

MedGAN was proposed by Armanious and colleagues [12] as a multi-purpose GAN for the transformation (translation, motion-correction, or denoising) of medical images. Unlike standard GANs, which transform noise into a desired image, MedGAN is conditional, taking as input a source image and transforming it into a desired target image. This type of operation is enabled by MedGAN's particular architecture and the losses that optimize it. Its generator is composed of several U-Net [22] architecture blocks that refine images as they propagate through a sequential encoder-decoder pathway. Modeled after PatchGAN, its discriminator decomposes generated images into 64 patches (each of size 16×16) in order to determine which are most likely to resemble a target image and which are least likely to do so. Together, the generator and discriminator are trained via adversarial loss, which places them in competition, respectively generating realistic images and identifying unrealistic images. While this leads the generator to output broadly sufficient images, they are often blurry. As a result, the generator uses three other loss functions that leverage feature extractors to produce more accurate images. Using the discriminator as a feature extractor, perceptual loss attempts to capture discrepancy between high frequency components of generated and target images, while content and style loss both use layers of VGG-19 [23] to quantify dissimilarities between the generated and target images. These loss formulations are given in the paper's Supplementary Materials. A schematic of our MedGAN denoising use-case and of the components of the MedGAN architecture are shown in Fig. 5, left panel.

3.2 MedGAN Methods

MedGAN Generator. The CasNet generator is built from U-Net modules [22], concatenated sequentially to create a richer output image. The blocks themselves are identically composed of an 8-layer encoding section followed by an 8-layer decoding section. The former employs convolutional layers with kernel sizes of four, stride length of 2, and convolutional filters of size 64, 128, 256, 512, 512, 512, 512, and 512. Given the input image size of $256 \times 256 \times 3$, these parameters eventually yield a $1 \times 1 \times 512$ object, which then feeds into the decoding section, built of 8 deconvolutional blocks with the same stride and kernel parameters as their convolutional counterparts and with filter sizes of 512, 1024, 1024, 1024, 1024, 512, 256, and 128. This section of the block deconvolves the 1×1 object

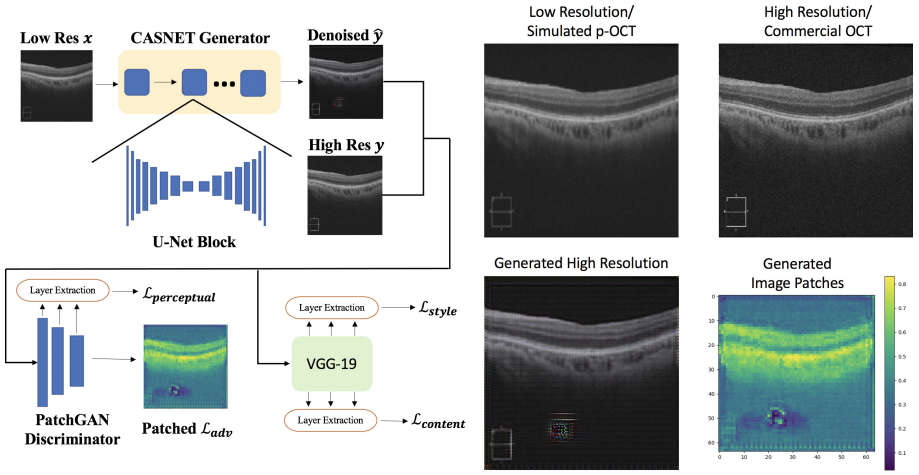

Fig. 5. Left Panel: Our MedGAN use-case is for image denoising; CasNet generator is composed of U-Net blocks; PatchGAN discriminator computes perceptual loss and patched adversarial loss; VGG-19 feature extractor computes style and content losses [12]. **Right Panel:** (Clockwise from top left) An example original low-resolution source image, high resolution target image, a heatmap showing the discriminator's prediction of image patches as real (yellow) or fake (blue), and the final MedGAN-generated image. The MedGAN-generated image is able to achieve a similar level of perceptual quality to the high resolution commercial OCT image, but the persistent artifact at the image's bottom affects this quality. Although these artifacts remained through many iterations of the MedGAN, this particular artifact is low-impact, away from important parts of the image (the retinal layers) and is well contained. (Color figure online)

with 512 channels back to a 256×256 object with 128 channels, allowing for the last deconvolutional layer to feed directly into the next block's first convolutional layer, ensuring modularity of the blocks. To output an image after the last U-Net block, we changed the output filter size of the last filter to 3 (to accommodate our RGB input images), such that a $256 \times 256 \times 3$ image results. Every convolutional and deconvolutional layer is followed by batch normalization and leaky ReLU layers, and the final deconvolutional layer employs a tanh activation function. We modified the original MedGAN architecture by using a sigmoid activation function at the output of the final block (to ensure output images remain in the range $[0, 1]$). Finally, encoding and decoding layers of the same dimension in each block are concatenated to ensure transfer of contextual information throughout the encoding-decoding pathway and to strengthen back-propagation.

MedGAN Discriminator. Unlike classic discriminators that output single values (the probability of the whole generated image being real or fake), MedGAN uses a patch discriminator that returns a 64×64 matrix of values indicating the probability that each 16×16 patch of an input image (either generated or target) is 'real' (each patch equals 1 for a target image, 0 for a generated

image.) This approach permits sharper results, especially in conjunction with non-adversarial losses like the perceptual loss. The architecture of this network is relatively straightforward, consisting of two convolutional layers with kernel sizes of 4, strides of 2, and spatial filters with parameters 64 in the first layer and 128 in the second. This ultimately produces a $64 \times 64 \times 128$ object, which can be compressed into a final $64 \times 64 \times 1$ matrix using a convolutional layer with its stride and spatial filter both set to 1, kernel size equal to 4, and a sigmoid activation function, as proposed in the original MedGAN design [12].

MedGAN Training. Training the MedGAN (via Lambda Labs Vector, RTX 3090 GPU) occurs in four iterations. In the first three, the generator produces images upon which losses can be computed and optimized. The discriminator receives no training during these iterations, only being optimized in the fourth and final iteration. This approach emphasizes the training of the generator over the discriminator, as the generator contains a much more complex architecture. This allows the two models to maintain a healthy training equilibrium, rather than creating a situation where the discriminator can dominate the generator.

3.3 MedGAN Results and Discussion

We chose a generator architecture with 6 CasNet blocks (as proposed in the MedGAN paper [12]), as it outperforms 1-block and 3-block architectures based on BRISQUE scores of generated outputs. We found the following loss hyperparameters to be optimal for the portable OCT data based on empirical tuning: Perceptual loss λ_{pi} of 1, Style loss λ_{sj} of 0.0001, and Content loss λ_{cj} of 0.0001. Smaller content and style losses may be preferred, because the VGG-19 feature extractor has been trained on ImageNet [24]; replacing VGG-19 with a network fine-tuned on OCT data may enhance quality of extracted features for p-OCT data, increasing the content/style loss contributions. We trained the model for 100 epochs (48 h) and with up to 1270 input images. Loss definitions and curves for all loss types are shown in the paper's Supplementary Materials. We found that the MedGAN-generated images (when scaled to match target image size) still exhibited significantly higher (poorer) average BRISQUE scores (44.6) than those of target commercial OCT images (35.5); in spite of visible reduction in noise (low perceptual loss) within the generated images, perceptual quality did not quantitatively match that of target images. As a more rigorous test, downstream AI performance using generated images is therefore shown via the blue ROC curve in Fig. 3 (left panel); **MedGAN generated** images exhibit an AUC of **0.774 (95% CI, 0.663–0.885)**, approaching that achieved by the target **commercial OCT** data of **0.8 (95% CI, 0.694–0.906)**. An example generated image and a heatmap indicating probability of 'realness' of patches within generated images according to the discriminator are also shown in Fig. 5, right panel.

We believe that mode collapse contributed to the persistent artifact occurring in the MedGAN output images. In typical GANs, this phenomenon transforms

all GAN inputs into a single output image, but mode collapse is not well-defined for conditional GANs like MedGAN. All of MedGAN's generated images appear to have a layer overlaying the original input image, with similar noise patterns and artifact locations in each. Therefore, while the conditional input remains different for all images, the generated artifacts are functionally identical.

4 Conclusions and Future Directions

In the first half of this paper, we used ESRGAN-based super-resolution *to enhance AI-based AMD detection using p-OCT images* and simulated paired training via the p-OCT point spread function. Perceptual quality of super-resolved outputs may be improved by first denoising the p-OCT data via existing high-performing, deep-learning based retinal-OCT denoising techniques [25,26] prior to ESRGAN super-resolution. An alternate future approach could involve using 'GAN-CIRCLE' and the cycle-consistency constraint to achieve super-resolution with unpaired training data and no task specific regularization [27]. In the second half of this paper, we sought to denoise/*improve perceptual quality* of simulated p-OCT data via MedGAN by reducing noise and perceptual loss (between source and target images). This resulted in AI performance close to that of target images, without significant quantitative perceptual quality improvement (lowering of BRISQUE scores). Artifacts in MedGAN-generated images could be eliminated by additional use of regularization and VGG-19 feature extractor fine-tuning on OCT data. Regularization strategies include adding instance noise to the MedGAN training input images and adding gradient penalties to the networks [28]. Another promising future direction is to integrate progressively growing generated images into GAN training [29], which would provide more stable image synthesis. Overall, we observed that GAN-based processing of p-OCT/simulated p-OCT data significantly improved AI-based AMD detection performance, in spite of not significantly changing perceptual quality, as assessed by classical metrics (BRISQUE) and the human eye. Our proof-of-principle findings stimulate future work toward AI-embedded p-OCT devices for eye disease detection, especially in situations when ophthalmic expertise or high quality testing data are not available.

References

1. Briganti, G., Le Moine, O.: Artificial intelligence in medicine: today and tomorrow. Front. Med. **7**, 27 (2020)
2. Moraru, A.D., Costin, D., Moraru, R.L., Branisteanu, D.C.: Artificial intelligence and deep learning in ophthalmology-present and future. Exp. Ther. Med. **20**(4), 3469–3473 (2020)
3. Teikari, P., Najjar, R.P., Schmetterer, L., Milea, D.: Embedded deep learning in ophthalmology: making ophthalmic imaging smarter. Ther. Adv. Ophthalmol. **11**, 2515841419827172 (2019)

4. Kim, S., Crose, M., Eldridge, W.J., Cox, B., Brown, W.J., Wax, A.: Design and implementation of a low-cost, portable OCT system. Biomed. Opt. Express **9**(3), 1232–1243 (2018)
5. Song, G., et al.: First clinical application of low-cost OCT. Transl. Vis. Sci. Technol. **8**(3), 61–61 (2019)
6. Tavakkoli, A., Kamran, S.A., Hossain, K.F., Zuckerbrod, S.L.: A novel deep learning conditional generative adversarial network for producing angiography images from retinal fundus photographs. Sci. Rep. **10**(1), 1–15 (2020)
7. Zheng, C., et al.: Assessment of generative adversarial networks model for synthetic optical coherence tomography images of retinal disorders. Transl. Vis. Sci. Technol. **9**(2), 29–29 (2020)
8. Lehtinen, J., et al.: Noise2Noise: learning image restoration without clean data. arXiv preprint arXiv:1803.04189 (2018)
9. Krull, A., Buchholz, T.O., Jug, F.: Noise2Void-learning denoising from single noisy images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2129–2137 (2019)
10. Wang, X., et al.: ESRGAN: enhanced super-resolution generative adversarial networks. In: Leal-Taixé, L., Roth, S. (eds.) ECCV 2018. LNCS, vol. 11133, pp. 63–79. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11021-5_5
11. 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)
12. Armanious, K., et al.: MedGAN: medical image translation using GANs. Comput. Med. Imaging Graph. **79**, 101684 (2020)
13. Thakoor, K., Bordbar, D., Yao, J., Moussa, O., Chen, R., Sajda, P.: Hybrid 3D–2D deep learning for detection of neovascularage-related macular degeneration using optical coherence tomography B-scans and angiography volumes. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), pp. 1600–1604 (2021)
14. Thakoor, K.A., et al.: A multimodal deep learning system to distinguish late stages of AMD and to compare expert vs. AI ocular biomarkers. Sci. Rep. **12**(1), 1–11 (2022)
15. Lugmayr, A., Danelljan, M., Timofte, R.: Unsupervised learning for real-world super-resolution. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), pp. 3408–3416 (2019)
16. Ledig, C., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4681–4690 (2017)
17. Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K.: Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 136–144 (2017)
18. Jolicoeur-Martineau, A.: The relativistic discriminator: a key element missing from standard GAN. arXiv preprint arXiv:1807.00734 (2018)
19. Wang, X., Yu, K., Chan, K.C.K., Dong, C., Loy, C.C.: BasicSR: open source image and video restoration toolbox. github.com/xinntao/BasicSR (2018)
20. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. **21**(12), 4695–4708 (2012)
21. Jang, H., McCormack, D., Tong, F.: Noise-trained deep neural networks effectively predict human vision and its neural responses to challenging images. PLoS Biol. **19**(12), e3001418 (2021)

22. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
23. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
24. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255 (2009)
25. Halupka, K.J., et al.: Retinal optical coherence tomography image enhancement via deep learning. Biomed. Opt. Express **9**(12), 6205–6221 (2018)
26. Qiu, B., et al.: Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function. Biomed. Opt. Express **11**(2), 817–830 (2020)
27. You, C., et al.: CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE). IEEE Trans. Med. Imaging **39**(1), 188–203 (2019)
28. Mescheder, L., Geiger, A., Nowozin, S.: Which training methods for GANs do actually converge? In: International Conference on Machine Learning, pp. 3481–3490. PMLR (2018)
29. Karras, T., Aila, T., Laine, S., Lehtinen, J.: Progressive growing of GANs for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196 (2017)