

ARTIFICIAL INTELLIGENCE

A review of generative adversarial network applications in optical coherence tomography image analysis



Jason Kugelman*, David Alonso-Caneiro, Scott A. Read, Michael J. Collins

Queensland University of Technology (QUT), Contact Lens and Visual Optics Laboratory, Centre for Vision and Eye Research, School of Optometry and Vision Science, Kelvin Grove, QLD 4059, Australia

KEYWORDS

Optical coherence tomography;
Image analysis;
Deep learning;
Generative adversarial networks

Abstract

Optical coherence tomography (OCT) has revolutionized ophthalmic clinical practice and research, as a result of the high-resolution images that the method is able to capture in a fast, non-invasive manner. Although clinicians can interpret OCT images qualitatively, the ability to quantitatively and automatically analyse these images represents a key goal for eye care by providing clinicians with immediate and relevant metrics to inform best clinical practice. The range of applications and methods to analyse OCT images is rich and rapidly expanding. With the advent of deep learning methods, the field has experienced significant progress with state-of-the-art-performance for several OCT image analysis tasks. Generative adversarial networks (GANs) represent a subfield of deep learning that allows for a range of novel applications not possible in most other deep learning methods, with the potential to provide more accurate and robust analyses. In this review, the progress in this field and clinical impact are reviewed and the potential future development of applications of GANs to OCT image processing are discussed.

© 2022 Spanish General Council of Optometry. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

The ability to non-invasively image the eye with optical coherence tomography (OCT) has transformed the fields of optometry and ophthalmology. Quantifying these images and numerically tracking changes (e.g. in tissue morphology) helps to better drive clinical decision making. OCT is a non-invasive coherent imaging method that is used to capture high-resolution cross-sectional images within optical scattering tissues such as the retina and choroid in the posterior segment of the eye (Fig. 1A). The analysis of such images is widely performed in both research and clinical practice for a

range of tasks including measuring retinal and choroidal layer thickness, for monitoring disease progression and for quantifying treatment outcomes (Fig. 1B).

Deep learning methods, a subfield of machine learning, have enhanced image analysis in OCT including for applications such as retinal and choroidal layer segmentation,¹⁻⁴ optic nerve head segmentation,² fluid region segmentation,^{3, 5} and pathology detection⁶⁻¹⁰ and/or grading.^{11,12} These methods can provide automatic, accurate and efficient solutions compared to traditional manual analysis by human observers. Generative adversarial networks (GANs) are an exciting new group of deep learning methods which have seen significant growth, progress and adoption in many areas in recent years, particularly in the field of medical image analysis where GANs have been applied to a number of ophthalmic imaging modalities

* Corresponding author.

E-mail address: j.kugelman@qut.edu.au (J. Kugelman).

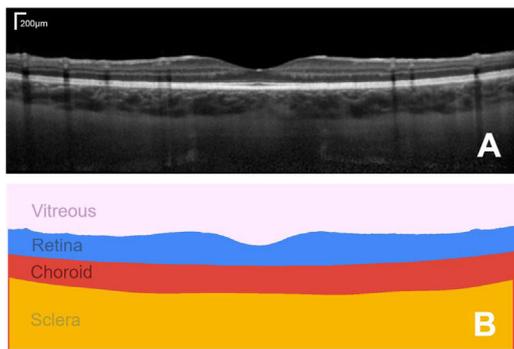


Fig. 1 Example foveal-centred spectral domain (SD) OCT scan of the posterior segment of the eye of a healthy subject, showing a cross-section of the retinal and choroidal tissue (A) and its corresponding segmented regions of interest, from which quantitative measures of tissue layer thickness can be derived (B). Bar in top right of A shows the scale.

including retinal fundus photography, OCT angiography (OCTA) and OCT. This review will focus exclusively on applications to OCT where GANs have demonstrated the ability to “generate” OCT images that are difficult to discern from real OCT images, even for experienced clinicians.¹³ For instance, Fig. 2 provides examples of four OCT images, three of which are generated by a GAN and one of which is a real OCT image (we encourage the reader to discern which it is). Given the high level of image quality, these GAN images may be used to train clinicians and even used as additional data for enhancing image analysis algorithms.

A GAN is a powerful and versatile deep learning method, particularly as a generative method for image synthesis (i.e. image generation) and related applications. The original GAN¹⁴ (Fig. 3) consists of two deep neural networks, a generator and a discriminator, competing against each other to improve their performance. Using a random noise vector, also called the latent code, as input, the generator’s goal is to synthesise realistic images such that the discriminator cannot distinguish them from the real images. In this context, the images produced by the generator are referred to as the “synthetic” or “fake” images. Through an iterative training process, the generator and discriminator both become stronger, with the generator learning to produce increasingly realistic images to “fool” the discriminator, while the discriminator also improves its ability to distinguish between real and fake images. The original GAN architecture,¹⁴ proposed in 2014, has been upgraded over the years to improve its performance, its training convergence and to expand its application capability. A few of the most

notable upgrades include that of Radford et al.¹⁵ who replaced the multi-layer perceptron (MLP) based architecture with a convolutional neural network (CNN) based one and proposed the deep convolutional GAN (DCGAN), which enabled the GAN to generate higher quality images. Mirza et al.¹⁶ extended the idea to the conditional GAN (cGAN) by incorporating an additional auxiliary input, such as a class label (e.g. type of disease in the image), which can be used to condition the input, thus providing control of the generated image (e.g. generate an image exhibiting particular disease characteristics). Expanding on the conditional approach, the discriminator in the auxiliary classifier GAN (ACGAN)¹⁷ was given the additional task of classifying each image (e.g. determining whether the image contains pathology A or pathology B). For exploiting both labelled (images with annotations) and unlabelled (images without annotations) data, the semi-supervised GAN¹⁸ operates in either supervised mode (classifying image into one of K classes) or unsupervised mode (real or fake), with the GAN discriminator possessing $K + 1$ outputs (an extra fake class is added). For example, taking the case of OCT disease classification, the GAN may be able to learn useful information (for the classification problem) about the images without these images being explicitly labelled with a disease (unlabelled). This can yield improved performance (more accurate classification) compared to using just labelled images and reduces the need for expert labelled data.

In this paper, a comprehensive review of GAN applications in optical coherence tomography is performed to identify where and how GANs have been used for OCT image analysis tasks, highlighting where possible its clinical application and motivation. Additionally, the gaps in the existing literature are discussed in terms of the potential for future work in the area where GANs may be used further to enhance OCT image analysis. This is a rapidly growing field and this review aims to provide an overview of the current state of development in GAN applications in OCT image analysis, and to stimulate future research within the area.

GAN applications

Given their demonstrated benefit for a number of applications in several areas of medical image analysis, GANs have been increasingly adopted for OCT image analysis tasks. Their most obvious and natural application is that of image synthesis (synthetic data generation). Hence, several studies have initially investigated the feasibility of generating realistic OCT images using GANs.^{13,19-21} However, there are a range of other applications for OCT image analysis that have been explored using GANs. These applications, which are summarised in the

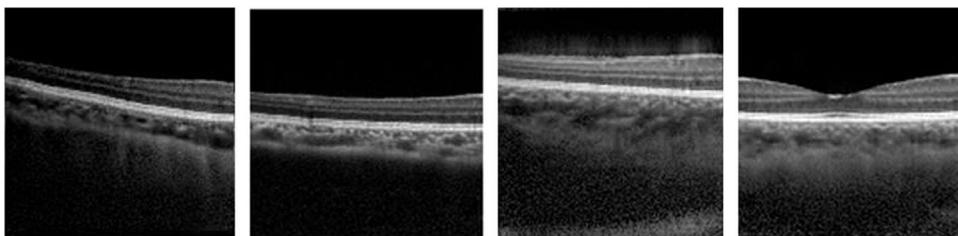


Fig. 2 Example OCT scans: three of which are generated by a GAN while one is real. We encourage the reader to discern which image is the real one (answer can be found in the Conclusion section).

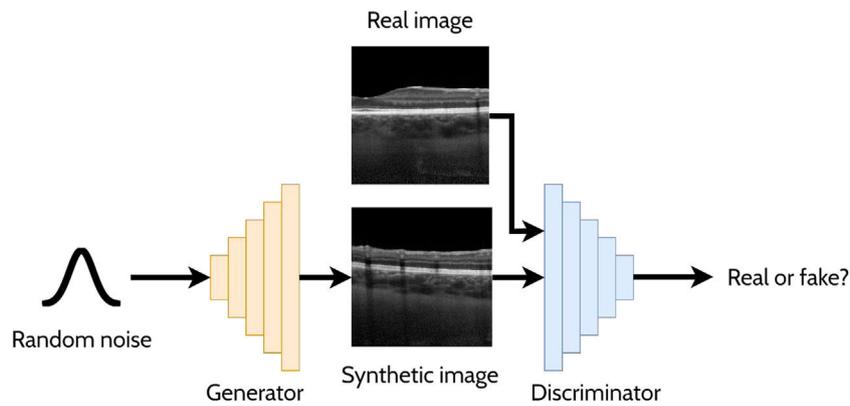


Fig. 3 Overview of a generative adversarial network (GAN) for generating cross-sectional retinal OCT images. The generator takes a random noise vector as input and produces a synthetic image. The discriminator is required to distinguish between the synthetic (fake) images (created by the generator) and the real images.

following sections, include: anomaly detection, domain translation, super-resolution, image enhancement, de-noising, data augmentation, shadow removal, segmentation, classification, and prediction. Although the main motivation of this work is to summarize and review these GAN applications, the clinical benefit and impact of these applications is also highlighted, where possible.

Anomaly detection

Anomaly detection refers to the task of detecting out-of-distribution (anomalous) samples (images) while anomaly localisation refers to the detection of anomalous regions with the samples (i.e. specific regions within images). For example, anomaly detection could involve the detection of OCT images of pathology from those that are healthy, whereas anomaly localisation could involve the localisation of pathology specific features (e.g. subretinal fluid) within these images. One approach for anomaly detection is by using a GAN to learn the data distribution and then using a second step to determine whether a given image belongs to the learnt GAN distribution (non-anomaly) or not (anomaly). This relies on the assumption that the GAN will only be able to generate images from within the learnt distribution and will not be able to reliably reconstruct anomalous samples. A handful of studies have explored anomaly detection and localisation using GANs in retinal OCT

images²²⁻²⁷ which will be summarised in the following sections. An illustrative example of using a GAN applied for anomaly detection is depicted in Fig. 4.

AnoGAN

The AnoGAN²² is initially trained to generate patches of normal OCT images. The method then adopts an iterative process to update a randomly initialised latent code to minimise the difference between the corresponding generated (reconstructed) OCT image patch (generated using the latent code) and the original target (real image) OCT image patch. Sufficiently large differences in the reconstructed image compared to the original are used to detect and flag whether the original image is anomalous. In this case, anomalies consisted of images that contained retinal fluid or hyperreflective foci. The specific spatial locations with the largest differences can be used to identify the locations of the anomalies (anomaly localisation). The updated version of this network, the f-AnoGAN²³ improves this process by incorporating an additional neural network (an encoder) to map the original target image to a latent code which is then provided as input to the generator. The reconstruction difference between the target image and the generated image can then be used to detect and localize anomalies in the OCT images in a similar way to the AnoGAN. This approach works under the assumption that normal

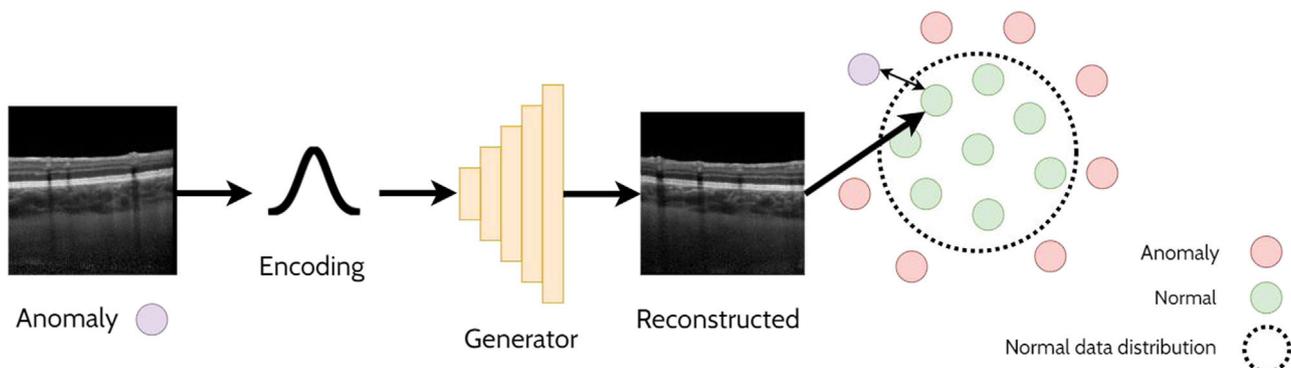


Fig. 4 Overview of an example anomaly detection application in OCT images using GANs.

OCT images will be correctly encoded and reconstructed by the GAN whereas out-of-distribution (anomalous) images will not, leading to a higher reconstruction error (difference between original and reconstructed images) as these images do not exist within the normal data distribution that was initially learnt by the GAN.

Other

Rather than calculating differences and detecting anomalies on the image level (e.g. AnoGAN), another approach uses a Sparse-GAN²⁴ which instead computes the anomalies with respect to the latent codes to detect lesions within OCT images including those of drusen, diabetic macular oedema (DME) and choroidal neovascularization (CNV). Wang et al.²⁶ used a weakly-supervised (using imprecise annotations) method based on CycleGANs²⁸ and reconstructed the normal anatomical structures from abnormal input images to detect and segment lesions, including drusen, DME and CNV. Zhou et al.²⁷ proposed the P-Net to perform anomaly detection by leveraging the relationship between structure (the regions and layers) and the texture of OCT images. Here, an image reconstruction module is proposed to 1) segment the image and obtain the semantic area mask, 2) encode the semantic area mask and image separately, 3) reconstruct the OCT image based off the two encodings. The difference in the extracted structure of the reconstructed OCT image and the original is used for anomaly detection.

Advantages and disadvantages

Compared to other machine learning-based anomaly detection methods, there are several advantages of GAN-based methods for anomaly detection in OCT. Firstly, these methods require no supervisory signal foregoing the need for any manual labelling compared to other deep learning approaches which may employ supervised, semi-supervised or weakly supervised techniques. Second, these methods are well suited to image data and yield higher accuracy and robustness given the ability of GANs to model complex data distributions and generate highly realistic and diverse images. Third, GAN-based anomaly detection methods require relatively little tuning and manual calibration and scale well with increasing data. However, one common drawback of GAN-based approaches for anomaly detection is the need for significant amounts of training data for the method to perform well. Nonetheless, GANs for anomaly detection have demonstrated significant potential to be useful tools in clinical practice for identifying and highlighting pathological regions within the retinal tissue.

Domain translation

Domain translation refers to the idea of taking a sample (image) from one domain and transferring that to its representation in another domain. For instance, this may involve translating an OCT image acquired by one OCT instrument (domain A) to resemble an image acquired by a second OCT instrument (domain B). Similarly, this could involve translating an image acquired using one set of scanning parameters

(domain A) to resemble an image acquired using a second set of scanning parameters (domain B). This has potential benefits including the ability to analyse images from multiple domains with an algorithm developed exclusively for one particular domain (e.g. images from a range of OCT instruments could be translated into a single ‘universal’ one for processing purposes). With these methods an important focus is ensuring that the ‘content’ or structure of a given image remains unchanged while just transferring the ‘style’ of the image. This is of particular importance to ensure such that clinically useful and reliable data can be derived. For OCT images, the content can refer to the structure of the tissue (i.e. critical anatomical features such as the position of the retinal layers and pathological regions) while the style can refer to features such as the particular level of speckle noise as well as the contrast between the retinal layers. A range of prior studies have investigated domain translation between two OCT instruments²⁹⁻³⁸ which will be summarised in the following sections. An example of applying a GAN to domain translation in OCT images is illustrated in Fig. 5.

CycleGAN based

A handful of methods have utilised CycleGAN as the basis of the domain translation approach. To improve retinal fluid segmentation and photoreceptor layer segmentation, Seebock et al.²⁹ and Romo-Bucheli et al.³⁰ used CycleGANs to translate between OCT images of two devices (Zeiss Cirrus to Heidelberg Spectralis images) taking patches of OCT images from a cross-domain dataset consisting of age-related macular degeneration (AMD), retinal vein occlusion (RVO) and DME. In another set of studies, Lazaridis et al.^{31,32} utilised an ensemble of spatially coherent CycleGANs to convert from time-domain OCT (TD-OCT) (low signal to noise ratio) to spectral domain OCT (SD-OCT) (higher signal to noise ratio), thus improving OCT image quality.

Other

He et al.³³ achieved a similar goal to the aforementioned CycleGAN methods but instead employed two discriminators (one for masks and one for boundaries) to constrain and ensure the correctness of the target OCT image anatomy. Yang et al.³⁴ adopted local and global discriminators and alignment modules for cross-device OCT lesion detection

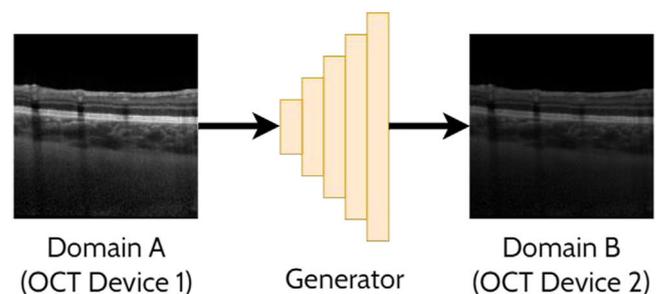


Fig. 5 Overview of an example domain translation application in OCT images using GANs.

including subretinal fluid (SRF), CNV and retinal pigment epithelium atrophy (RPEA). Wang et al.³⁵ used a domain adaptation model to improve classification accuracy for retinopathy detection from cross-domain OCT images by extracting invariant and discriminative characteristics shared by the images from two different OCT instruments (Heidelberg Spectralis and Zeiss Cirrus). Zhang et al.³⁶ ensured that the noise patterns were similar between domains, and used a noise adaptation approach consisting of two discriminators, one to ensure that the noise patterns were similar to the target domain and the other to enforce the content (structure of the tissue and layers) to be preserved in the generated OCT images. Bian et al.³⁷ used adversarial learning (i.e. using a GAN) for unsupervised domain adaptation with two OCT instruments (Heidelberg Spectralis and Optovue) as part of the uncertainty-aware domain alignment method for retinal and choroidal layer segmentation in OCT images. Chai et al.³⁸ incorporated a perceptual loss³⁹ on the output segmentation maps such that the choroidal structure is preserved between different OCT domains in an unsupervised manner.

Advantages and disadvantages

A clear advantage of using GAN-based approaches for domain translation is the high level of realism and quality of the generated images. Additionally, the rules of the transformation are automatically learnt and do not need to be handcrafted or manually specified. There is then less reliance on domain-specific knowledge and understanding of the domains in question which can allow for more accessible and efficient development of these methods. As highlighted above, GANs are also adept at learning the transformation between domains without requiring labelled pairs of images. This is a significant advantage over other methods as it means that these models are more readily trainable and do not require the time-consuming and often tedious collection of additional paired data, which can also be difficult to obtain depending on the application and type of data required. Despite the advantages, one disadvantage of GAN-based approaches for domain translation is the need for large and diverse datasets for training which can be difficult to acquire in certain situations, particularly in the medical domain. This problem is amplified for domain translation in particular, as sufficient data from two domains are required, rather than just one. As discussed, GANs have been applied

in numerous studies to translate images from one domain to another such as translating an image from the style of one instrument to another. However, there is no study which has investigated a single model for translating between multiple pairs of instruments which would allow for a more unified and efficient solution.

Super-resolution and de-noising

Super-resolution and de-noising can both be considered as specific examples of domain translation. Super-resolution involves taking a low-resolution image and transforming this to its high-resolution equivalent which may provide additional information or better visualization of the ocular tissue. With the goal of improving the analysis of noisy OCT images, de-noising involves the process of reducing the speckle noise which is an inherent property of coherent imaging techniques such as OCT. One approach for reducing speckle noise is to perform real-time image averaging during scan acquisition with the OCT instrument. However, this requires a prolonged scan acquisition time which is not always feasible, particularly in the cases of poor fixation. Hence the need for post-processing methods such as a GAN to perform this task and enable improved visualisation of the tissue structures in the image which can prove challenging in the presence of high levels of noise. Examples of applying GANs for super-resolution and de-noising in OCT images are depicted in Fig. 6.

Summary

There are several existing methods for super-resolution⁴⁰⁻⁴⁶ and de-noising^{40,47-61} for OCT images. Common trends include the use of a perceptual (content) loss (using a pre-trained network) and/or a structural similarity loss for preserving structural information within the generated OCT images,^{40,41,45,47,55-59} and the adoption of PatchGAN discriminators (discriminate on the scale of patches of the images rather than the full images).^{42,43,48,53,54,60} There are several examples of both supervised (i.e. using paired data with both the noisy and averaged image pairs)^{41,43,49,51,54-56,59} and unsupervised (i.e. using unpaired data with only the noisy images)^{42,50,52-54} approaches with just a single semi-supervised (using both paired and unpaired data) approach published to date.⁵⁸ The following sections summarise these studies categorised into supervised and unsupervised methods.

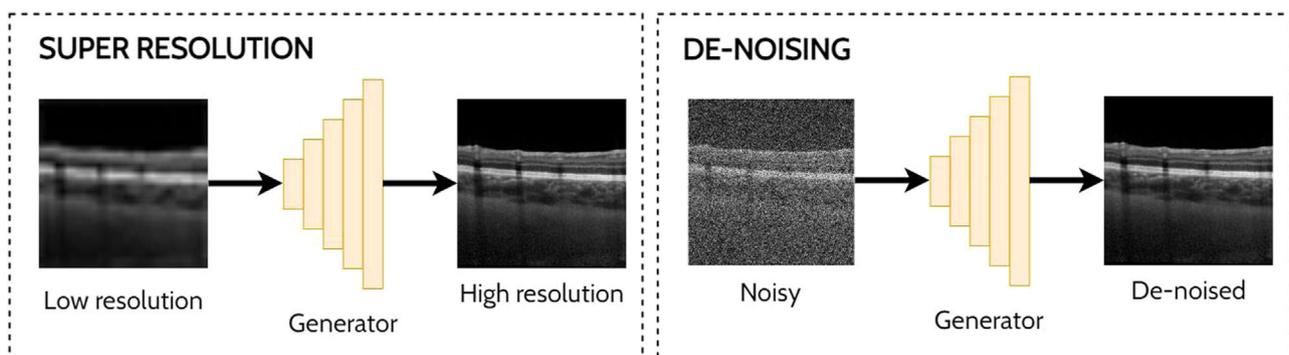


Fig. 6 Overview of an example super-resolution application (left) and de-noising application (right) in OCT images using GANs.

Supervised methods

For supervised methods, the target (reference) super-resolved or de-noised OCT images are available and explicitly utilised. Huang et al.⁴⁰ performed simultaneous de-noising and super-resolution of AMD OCT images and proposed the SDR-OCT method with a network based on the dense deep back-projection network for super-resolution.⁶² Hao et al.⁴³ used a Pix2Pix GAN to reconstruct high-bit depth OCT scans (12-bit) from low ones (3–8 bit). The OCT-GAN⁴⁷ aimed to remove noise from optic nerve head images and uses three pre-trained networks to preserve the content and style of the OCT images, as well as to avoid blurring. Ma et al.⁴⁸ used a supervised image-to-image translation approach for de-noising OCT images of patients with central serous chorioretinopathy (CSC) and pathological myopia (PM) and combined an edge loss function to increase sensitivity to and better preserve edge based details (i.e. the retinal layer boundaries). Chen et al.⁴⁹ proposed the DN-GAN for SD-OCT de-noising in optic nerve head images (Heidelberg Spectralis) using an image-to-image translation approach (convert high noise to low noise) which used two reconstruction losses, one in the spatial domain (as standard) and the other in the frequency domain. For both normal and pathological eyes, Guo et al.⁵⁷ also utilised CycleGAN for unpaired SD-OCT (Heidelberg Spectralis) de-noising proposing the structure-aware noise reduction GAN (SNR-GAN) utilising a structure-aware loss (SSIM) to aid in preserving the different structures within the OCT images during the de-noising process. Viedma et al.,⁶¹ utilised CycleGANs for de-noising healthy SD-OCT images with the goal of improving semantic segmentation performance of the retinal layers. The SiameseGAN⁵⁵ uses an additional siamese twin network taking matching (identical ground truth images) and non-matching (ground truth and generated images) pairs to improve GAN training and produce higher-quality de-noised SD-OCT (Bioptigen) images of both normal and AMD patients.

Unsupervised methods

For unsupervised methods, the target OCT scans (i.e. de-noised or super-resolved image) are not available for training while only some are available for semi-supervised methods. Das et al.⁴² used CycleGANs for unsupervised OCT super-resolution to improve AMD diagnosis. For both normal and AMD patients, Guo et al.⁵⁰ performed unsupervised de-noising of SD-OCT (Bioptigen) image patches without reference clean images, instead assuming the speckle noise distribution using ‘background’ regions from the noisy OCT images. To de-noise retinal OCT (Heidelberg Spectralis) images of various diseases (and different stages of disease), Manakov et al.⁵² used a modified CycleGAN (HDCycleGAN) in an unsupervised fashion with just a single discriminator such that the discriminator can focus on the differences between the noisy and the averaged images. Using a CycleGAN approach with two separate encoders, Huang et al.^{53,54} proposed that noisy SD-OCT images (from both normal and AMD patients) can be represented in both a content space and a noise space while the clean images are represented solely in the content space (no noise). Wang et al.⁵⁸ performed semi-

supervised de-noising of swept-source OCT (SS-OCT) images (acquired using Topcon Atlantis) using a capsule network⁶³ based conditional GAN again employing a structural similarity loss to preserve important structural information (i.e. the retinal layers). Here, the model is initially trained on the paired images (supervised). Afterwards, semi-supervised training is performed with both supervised (reconstruction, structural similarity, and adversarial losses) and unsupervised (adversarial loss only) modes which exploits the use of unpaired data to improve de-noising performance, therefore requiring less data to be acquired in the first place to train the model.

Advantages and disadvantages

There is extensive work for image enhancement techniques of OCT images using GANs, predominantly for de-noising and super-resolution. One particular advantage of GANs for image enhancement and de-noising is that these can be used without paired samples. For instance, it may be difficult to acquire paired low-resolution and high-resolution images that are aligned precisely at the same location within the tissue. Another advantage of using GANs compared to other methods in these cases is the ability to synthesise images that are highly realistic. This is a particular challenge with other image enhancement methods in general which often produce images which contain artifacts or overly blurred features. Indeed, retaining the anatomical structure and details within the images is of utmost importance such that clinically derived metrics are unaffected by the enhancement process. However, a particular ongoing challenge in developing image enhancement methods is determining the best metrics to use to evaluate the quality of the enhanced images. Despite their advantages, one particular disadvantage of GANs is the need for significant amounts of data to train the models well. Without this, GANs can fail to generalise well to new, unseen data. This can lead to erroneous images or even the case where the translated, enhanced images can resemble those in the training set, rather than resembling the input image of interest.

Data augmentation

Deep learning methods generally require sufficiently large and diverse annotated datasets for ample performance. However, this is often difficult, particularly in the medical domain, due to privacy concerns and the issues posed by the significant time and cost to collect and annotate large volumes of data. One method to address this is to employ data augmentation to artificially enhance the dataset using simple image transformations such as rotations and contrast adjustments. Data augmentation is another application of GANs which involves combining the synthetic data generated by a GAN with the original real training data. In general, this increase in data can help the training and generalization of the model. The additional benefit here is that the GAN can synthesise new samples from within the learnt distribution which are beneficial for the learning task of interest. For instance, if we consider the task of OCT retinal layer segmentation with a low number of images to train the model, a GAN can be used to effectively increase the number of

training images (i.e. generate new ‘synthetic’ samples). Hence, the combination of the new synthetic training data and the real training data can potentially improve the performance of a particular task compared to just using real training data alone. GAN-based data augmentation has been used to enhance OCT retinal segmentation,⁶⁴⁻⁶⁸ and classification^{69,70} methods which are summarised in the following sections. Fig. 7 provides an illustrative example of an OCT data augmentation application using GANs.

Segmentation

For segmentation, Kugelmann et al.^{64,66} performed data augmentation using conditional GANs to enhance the training set of SD-OCT (Heidelberg Spectralis) patches for retinal and choroidal layer segmentation in images of healthy eyes. Mahapatra et al.⁶⁵ performed data augmentation for retinal OCT segmentation of pathological regions including intraretinal fluid (IRF), SRF, and pigment epithelial detachment (PED) in images of patients presenting with AMD and RVO. SD-OCT scans were acquired from OCT instruments from three different vendors: Cirrus HD-OCT, Spectralis and Topcon. Here, they adopted a geometry-aware GAN (GeoGAN) to generate a new image and segmentation mask by jointly encoding an image and mask and performing hierarchical uncertainty sampling within the generator.

Classification

For data augmentation in classification, Chen et al.⁶⁹ used conditional GANs to generate diverse images of different modalities including for OCT images of normal patients and those with retinal diseases (CNV, DME, drusen) to be used for retinal disease detection. Diversity is introduced by inserting noise to the middle layers of the generator to introduce variations to the images. Given the unbalanced nature of the data, Yoo et al.⁷⁰ used a CycleGAN to generate pathological OCT images (from healthy ones) which were then used to augment data to train a classifier of rare retinal diseases, such as CSC, macular telangiectasia, Stargardt disease, retinitis pigmentosa, and macular hole as well as more prevalent retinal diseases such as DME, drusen and CNV. GANs were trained individually for each of the diseases.

Advantages and disadvantages

There are several advantages of GANs for synthesising images for data augmentation including the high level of quality and diversity possible with state-of-the-art GAN methods.

Additionally, GANs automatically learn the data distribution in question, removing the need for manually specifying the rules for augmenting the images and can often be easily combined with standard data augmentation techniques to improve performance. However, one significant limitation of GANs for data augmentation is that they become restricted to only generate images that are similar to the training set which can limit their effectiveness depending on the application and dataset. Indeed, images that do not resemble those within the training set will be flagged as fake by the discriminator and the generator will therefore not tend to produce those. Overall, there are just a few studies, however given these results highlight a real potential for this application in OCT, further studies are warranted in the future. Utilising synthetic data for data augmentation has demonstrated a use for improving the accuracy of classification and segmentation methods in OCT particularly in cases where large and diverse labelled datasets are difficult to obtain.

Other applications

There are several other applications of GANs to OCT image analysis including a handful of studies investigating shadow removal,^{47,71} segmentation,⁷²⁻⁷⁴ classification⁷⁵ and prediction.⁷⁶ De-shadowing OCT images involves the removal of shadow artifacts caused by the obstruction and optical scattering by the retinal blood vessels. Such shadows can be problematic for image analysis due to the difficulty in visualising the structure of the tissue in the shadowed regions. DeshadowGAN,⁷¹ and OCT-GAN⁴⁷ both adopt supervised approaches employing shadow removal and shadow detection networks. The DeshadowGAN shadow detection network combines a content loss (to ensure non-shadow regions of the images are not altered), style loss (to ensure image texture remained similar), total variation loss (to avoid checkerboard artifacts) and shadow loss (based on the shadow detector network, to ensure shadows had been removed). Ouyang et al.⁷² used conditional GANs to perform pre-segmentation of anterior OCT images (both corneal and limbal images) to remove speckle noise and specular artifacts above the shallowest interface. Liu et al.⁷³ used an image-to-image translation approach with a GAN to perform semi-supervised segmentation of retinal layers and fluid regions in OCT images of DME. Here, they leverage the unlabelled data to improve performance by using a fully-convolutional discriminator with a confidence map output to highlight trustworthy segmented regions in the unannotated images which are then used to guide the training. Das et al.⁷⁵ applied a semi-supervised approach for classification of

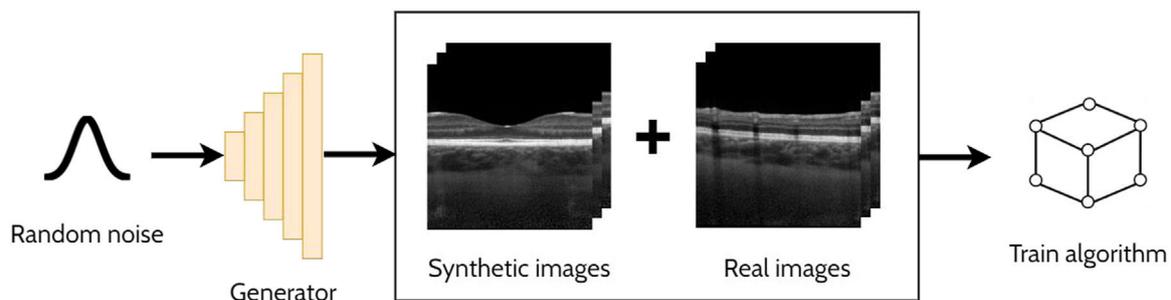


Fig. 7 Overview of an example application performing data augmentation using a GAN.

normal OCT images and those exhibiting ocular pathologies (AMD, DME), simultaneously training for classification (labelled images) and in an adversarial manner (unlabelled images). For prediction, Lee et al.⁷⁶ used GANs to predict OCT scans post-treatment for age related macular degeneration. Fig. 8 provides illustrative examples of segmentation, shadow removal and classification for OCT images using GANs.

Conclusions and future work

This review has covered the extensive use of GANs for applications in OCT image analysis including for anomaly detection, domain translation, super-resolution, de-noising, data augmentation, segmentation, classification and prediction. In many cases, GANs provide an improved approach compared to other deep learning and traditional algorithms and techniques. Indeed, for many of these applications GANs are state-of-the-art in terms of their accuracy (e.g. anomaly detection) and the quality of the output images (e.g. de-noising). The innate ability of GANs to model and learn complex data distributions is particularly suited to image-based tasks, hence their state-of-the-art performance and increasing use in medical image analysis including OCT. In OCT, the power of GANs is also highlighted by their demonstrated ability to synthesise realistic images that can fool even experienced clinicians.¹³ Fig. 2 illustrated the difficulty in discerning real OCT scans from fake ones generated by a GAN (in this case only the second leftmost image is in fact real).

While some applications have received more attention than others there are also some notable gaps. Despite the studies in domain translation with two OCT instruments, no studies have investigated this application in the multi-domain setting with more than two OCT instruments. A single model capable of translating images between a range of

different OCT instruments would allow for more unified and streamlined approaches to OCT image analysis and would allow algorithms developed using particular imaging modalities to be compatible with other modalities. There is also limited work performing adversarial segmentation of posterior OCT images (either supervised or semi-supervised). In the supervised setting there is the potential for improving the segmentation accuracy of existing approaches by incorporating a GAN loss while in the semi-supervised setting unlabelled data can be exploited to improve performance which is of significant interest in medical image analysis where labelling large datasets can be costly, time consuming or infeasible. There is also very limited work for patch-based data augmentation for OCT segmentation with only a single set of studies on healthy data and similarly limited studies investigating data augmentation using GANs for existing OCT semantic segmentation methods. Fast and accurate layer segmentation, especially in a low-data setting, is crucial in both research and clinical practice.

Although this review has focused on OCT in a comprehensive manner, there are similar GAN applications in other areas of ophthalmic image analysis including for retinal fundus photography (RF)⁷⁷⁻⁸³ and OCT angiography (OCTA).⁸⁴ The area of GANs is rapidly evolving and this review aims to provide a window into their wide and increasing level of application to optical coherence tomography. It is hoped that this review will aid future research direction and projects in this research space. GANs have shown promise in various OCT image analysis applications and are likely to further contribute to this field in the future, particularly considering the rapid advances that have been observed in the short time since GANs have been introduced. As a consequence, research and clinical decision making should positively benefit from these automatic and reliable image analysis methods.

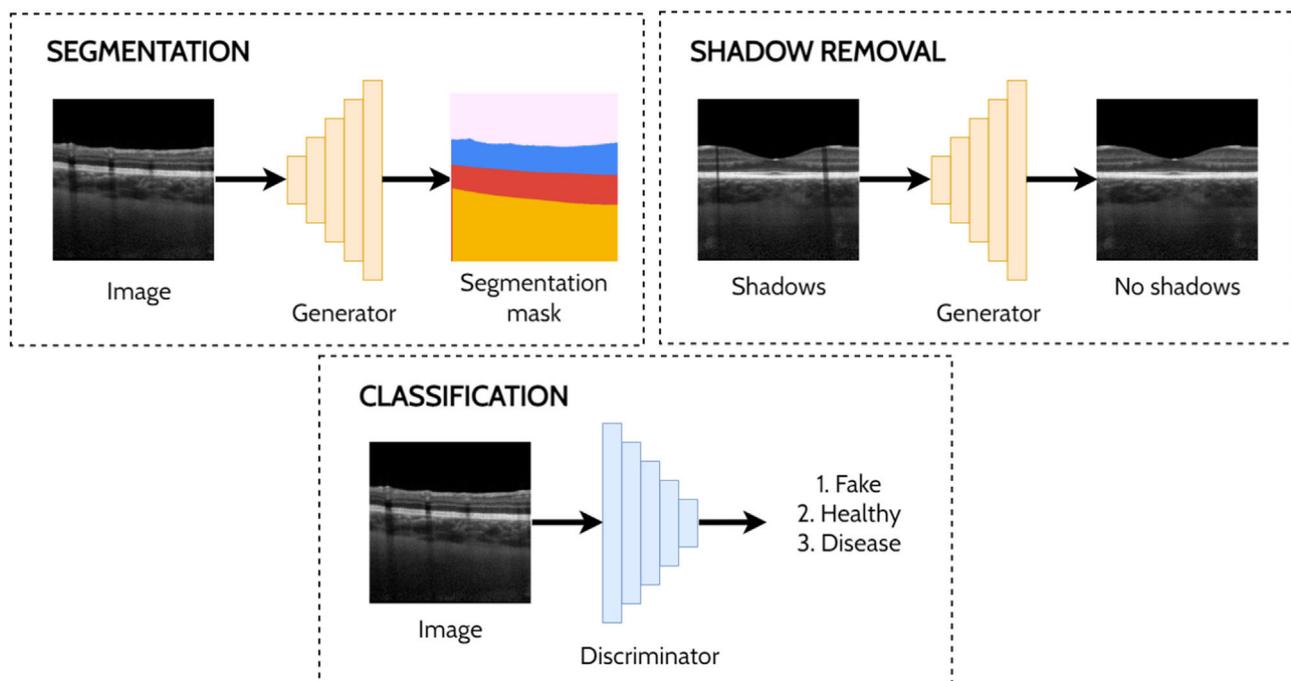


Fig. 8 Overview of example segmentation, shadow removal, and classification applications using GANs.

Competing interests

The authors declare no competing interests.

References

- Kugelman J, Alonso-Caneiro D, Read SA, et al. Automatic choroidal segmentation in OCT images using supervised deep learning methods. *Sci Rep*. 2019;9:13298.
- Devalia SK, Renukanand PK, Sreedhar BK, et al. DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. *Biomed Opt Express*. 2018;9:3244–3265.
- Roy AG, Conjeti S, Karri SPK, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed Opt Express*. 2017;8:3627–3642.
- Venhuizen FG, van Ginneken B, Liefers B, et al. Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks. *Biomed Opt Express*. 2017;8:3292–3316.
- Chen M, Ma W, Shi L, Li M, Wang C, Zheng G. Multiscale dual attention mechanism for fluid segmentation of optical coherence tomography images. *Appl Opt*. 2021;60:6761–6768.
- Lee CS, Baughman DM, Lee AY. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. *Ophthalmology Retina*. 2017;1:322–327.
- Yoon J, Han J, Park JI, et al. Optical coherence tomography-based deep-learning model for detecting central serous chorioretinopathy. *Sci Rep*. 2020;10:18852.
- Awais M, Müller H, Tang TB, Meriaudeau F. Classification of sd-oct images using a deep learning approach. 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA): IEEE. 2017:489–492.
- Wang D, Wang L. On OCT Image Classification via Deep Learning. *IEEE Photonics J*. 2019;11:1–14.
- Lu W, Tong Y, Yu Y, Xing Y, Chen C, Shen Y. Deep learning-based automated classification of multi-categorical abnormalities from optical coherence tomography images. *Transl Vis Sci Technol*. 2018;7:41–41.
- Sandhu HS, Eltanboly A, Shalaby A, Keynton RS, Schaal S, El-Baz A. Automated diagnosis and grading of diabetic retinopathy using optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2018;59:3155–3160.
- Raja H, Hassan T, Akram MU, Werghi N. Clinically verified hybrid deep learning system for retinal ganglion cells aware grading of glaucomatous progression. *IEEE Trans Biomed Eng*. 2021;68:2140–2151.
- Zheng C, Xie X, Zhou K, et al. Assessment of generative adversarial networks model for synthetic optical coherence tomography images of retinal disorders. *Transl Vis Sci Technol*. 2020;9:29–29.
- Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. Montreal, Canada: MIT Press. 2014:2672–2680.
- Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv:1511.06434*. 2015.
- Mirza M., Osindero S. Conditional generative adversarial nets. *arXiv:1411.1784*. 2014.
- Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier gans. International conference on machine learning: PMLR. 2017:2642–2651.
- Odena A. Semi-Supervised Learning with Generative Adversarial Networks. *arXiv:1606.01583 [stat.ML]*. 2016.
- Zheng C, Bian F, Li L, et al. Assessment of generative adversarial networks for synthetic anterior segment optical coherence tomography images in closed-angle detection. *Transl Vis Sci Technol*. 2021;10:34–34.
- Odaibo S.G. Generative adversarial networks synthesize realistic OCT images of the retina. *arXiv:1902.06676 [cs.CV]*. 2019.
- Zha X, Shi F, Ma Y, Zhu W, Chen X. Generation of Retinal OCT Images with Diseases Based on cGAN. *SPIE*. 2019.
- Schlegl T, Seeböck P, Waldstein SM, Schmidt-Erfurth U, Langs G. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. *Inf Process Med Imaging*. 2017:146–157.
- Schlegl T, Seeböck P, Waldstein SM, Langs G, Schmidt-Erfurth U. f-AnoGAN: fast unsupervised anomaly detection with generative adversarial networks. *Med Image Anal*. 2019;54:30–44.
- Zhou K, Gao S, Cheng J, et al. Sparse-Gan: sparsity-constrained generative adversarial network for anomaly detection in retinal OCT image. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). 2020:1227–1231.
- Zhang C, Wang Y, Zhao X, et al. Memory-augmented anomaly generative adversarial network for retinal OCT images screening. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). 2020:1971–1974.
- Wang J, Li W, Chen Y, et al. Weakly supervised anomaly segmentation in retinal OCT images using an adversarial learning approach. *Biomed Opt Express*. 2021;12:4713–4729.
- Zhou K, Xiao Y, Yang J, et al. Encoding structure-texture relation with P-Net for anomaly detection in retinal images. Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part 16: Springer. 2020:360–377.
- Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. 2017:2223–2232.
- Seeböck P, Romo-Bucheli D, Waldstein S, et al. Using cyclegans for effectively reducing image variability across oct devices and improving retinal fluid segmentation. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). 2019:605–609.
- Romo-Bucheli D, Seeböck P, Orlando JI, et al. Reducing image variability across OCT devices with unsupervised unpaired learning for improved segmentation of retina. *Biomed Opt Express*. 2020;11:346–363.
- Lazaridis G, Lorenzi M, Ourselin S, Garway-Heath D, et al. Enhancing OCT signal by fusion of GANs: improving statistical power of glaucoma clinical trials. In: Shen D, Liu T, Peters TM, eds. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. Springer, Cham; 2019:3–11.
- Lazaridis G, Lorenzi M, Ourselin S, Garway-Heath D. Improving statistical power of glaucoma clinical trials using an ensemble of cyclical generative adversarial networks. *Med Image Anal*. 2021;68:101906.
- He Y, Carass A, Liu Y, Saidha S, Calabresi P, Prince J. Adversarial domain adaptation for multi-device retinal OCT segmentation. *SPIE Medical Imaging*. 2020;113,13.
- Yang S, Zhou X, Wang J, et al. Unsupervised domain adaptation for cross-device OCT lesion detection via learning adaptive features. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). 2020:1570–1573.
- Wang J, Chen Y, Li W, et al. Domain adaptation model for retinopathy detection from cross-domain OCT images. In: Tal A, Ismail Ben A, Marleen de B, Maxime D, Herve L, Christopher P, eds. *Proceedings of the Third Conference on Medical Imaging with Deep Learning*. 2020:795–810. Vol 121.
- Zhang T, Cheng J, Fu H, et al. Noise adaptation generative adversarial network for medical image analysis. *IEEE Trans Med Imaging*. 2020;39:1149–1159.

37. Bian C, Yuan C, Wang J, et al. Uncertainty-aware domain alignment for anatomical structure segmentation. *Med Image Anal.* 2020;64: 101732.
38. Chai Z, Zhou K, Yang J, et al. Perceptual-assisted adversarial adaptation for choroid segmentation in optical coherence tomography. *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. 2020:1966–1970.
39. Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. *European Conference On Computer Vision*. Springer; 2016:694–711.
40. Huang Y, Lu Z, Shao Z, et al. Simultaneous denoising and super-resolution of optical coherence tomography images based on generative adversarial network. *Opt Express.* 2019;27: 12289–12307.
41. Yuan Z, Yang D, Pan H, Liang Y. Axial super-resolution study for optical coherence tomography images via deep learning. *IEEE Access.* 2020;8:204941–204950.
42. Das V, Dandapat S, Bora PK. Unsupervised super-resolution of OCT images using generative adversarial network for improved age-related macular degeneration diagnosis. *IEEE Sens J.* 2020;20:8746–8756.
43. Hao Q, Zhou K, Yang J, et al. High signal-to-noise ratio reconstruction of low bit-depth optical coherence tomography using deep learning. *J Biomed Opt.* 2020: 123702.
44. Pan H, Yang D, Yuan Z, Liang Y. More realistic low-resolution OCT image generation approach for training deep neural networks. *OSA Continuum.* 2020;3:3197–3205.
45. Cao S, Yao X, Koirala N, et al. Super-resolution technology to simultaneously improve optical & digital resolution of optical coherence tomography via deep learning. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2020:1879–1882.
46. Jeihouni P, Dehzangi O, Amireskandari A, Rezai A, Nasrabadi NM. Gan-based super-resolution and segmentation of retinal layers in optical coherence tomography scans. *2021 IEEE International Conference on Image Processing (ICIP)*. 2021:46–50.
47. Cheong H, Krishna Devalla S, Chuangsuwanich T, et al. OCT-GAN: single step shadow and noise removal from optical coherence tomography images of the human optic nerve head. *Biomed Opt Express.* 2021;12:1482–1498.
48. Ma Y, Chen X, Zhu W, Cheng X, Xiang D, Shi F. Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN. *Biomed Opt Express.* 2018;9: 5129–5146.
49. Chen Z, Zeng Z, Shen H, Zheng X, Dai P, Ouyang P. DN-GAN: denoising generative adversarial networks for speckle noise reduction in optical coherence tomography images. *Biomed Signal Process Control.* 2020;55: 101632.
50. Guo A, Fang L, Qi M, Li S. Unsupervised denoising of optical coherence tomography images with nonlocal-generative adversarial network. *IEEE Trans Instrum Meas.* 2021;70:1–12.
51. Hasan MJ, Alom MS. *Deep Learning Based Retinal OCT Image Denoising With Generative Adversarial Network*. Basherhat, Bangladesh: Department of Electrical and Electronic Engineering, Hajee Mohammad Danesh Science and Technology University; 2020.
52. Manakov I, Rohm M, Kern C, Schworm B, Kortuem K, Tresp V, et al. Noise as domain shift: denoising medical images by unpaired image translation. In: Wang Q, Milletari F, Nguyen HV, eds. *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*. 2019:3–10.
53. Huang Y, Xia W, Lu Z, et al. Disentanglement network for unsupervised speckle reduction of optical coherence tomography images. In: Martel AL, Abolmaesumi P, Stoyanov D, eds. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. 2020:675–684.
54. Huang Y, Xia W, Lu Z, et al. Noise-powered disentangled representation for unsupervised speckle reduction of optical coherence tomography images. *IEEE Trans Med Imaging.* 2020.
55. Kande NA, Dakhane R, Dukkupati A, Yalavarthy PK. SiameseGAN: a generative model for denoising of spectral domain optical coherence tomography images. *IEEE Trans Med Imaging.* 2021;40: 180–192.
56. Dong Z, Liu G, Ni G, Jerwick J, Duan L, Zhou C. Optical coherence tomography image denoising using a generative adversarial network with speckle modulation. *J Biophotonics.* 2020;13: e201960135.
57. Guo Y, Wang K., Yang S., et al. Structure-aware noise reduction generative adversarial network for optical coherence tomography image. In: Fu H, Garvin MK, MacGillivray T, Xu Y, Zheng Y, eds. *Ophthalmic Medical Image Analysis*. 2019:9–17.
58. Wang M, Zhu W, Yu K, et al. Semi-supervised capsule cGAN for speckle noise reduction in retinal OCT images. *IEEE Trans Med Imaging.* 2021;40:1168–1183.
59. Halupka KJ, Antony BJ, Lee MH, et al. Retinal optical coherence tomography image enhancement via deep learning. *Biomed Opt Express.* 2018;9:6205–6221.
60. Zhou Y, Yu K, Wang M, et al. Speckle noise reduction for OCT images based on image style transfer and conditional GAN. *IEEE J Biomed Health Inform.* 2022;26:139–150.
61. Viedma IA, Alonso-Caneiro D, Read SA, Collins MJ. OCT retinal image-to-image translation: analysing the use of CycleGAN to improve retinal boundary semantic segmentation. *2021 Digital Image Computing: Techniques and Applications (DICTA)* 2021:01–08.
62. Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018:1664–1673.
63. Sabour S., Frosst N., Hinton G.E. Dynamic routing between capsules. *arXiv:1710.09829 [cs.CV]*. 2017.
64. Kugelman J, Alonso-Caneiro D, Read SA, Vincent SJ, Chen FK, Collins MJ. Data augmentation for patch-based OCT chorio-retinal segmentation using generative adversarial networks. *Neural Comput Appl.* 2021.
65. Mahapatra D, Bozorgtabar B, Shao L. Pathological retinal region segmentation from OCT images using geometric relation based augmentation. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020:9608–9617.
66. Kugelman J, Alonso-Caneiro D, Read SA, Vincent SJ, Chen FK, Collins MJ. Constructing synthetic chorio-retinal patches using generative adversarial networks. *2019 Digital Image Computing: Techniques and Applications (DICTA)*. 2019:1–8.
67. Kugelman J, Alonso-Caneiro D, Read SA, Vincent SJ, Chen FK, Collins MJ. Dual image and mask synthesis with GANs for semantic segmentation in optical coherence tomography. *2020 Digital Image Computing: Techniques and Applications (DICTA)*. 2020:1–8.
68. Wu M, Cai X, Chen Q, et al. Geographic atrophy segmentation in SD-OCT images using synthesized fundus autofluorescence imaging. *Comput Methods Programs Biomed.* 2019;182: 105101.
69. Chen X., Li Y., Yao L., Adeli E., Zhang Y. Generative Adversarial U-Net for Domain-free Medical Image Augmentation. *arXiv:2101.04793 [eess.IV]*. 2021.
70. Yoo TK, Choi JY, Kim HK. Feasibility study to improve deep learning in OCT diagnosis of rare retinal diseases with few-shot classification. *Med Biol Eng Comput.* 2021;59:401–415.
71. Cheong H, Devalla SK, Pham TH, et al. DeshadowGAN: a deep learning approach to remove shadows from optical coherence tomography images. *Transl Vis Sci Technol.* 2020;9: 23-23.
72. Ouyang J, Mathai TS, Lathrop K, Galeotti J. Accurate tissue interface segmentation via adversarial pre-segmentation of anterior segment OCT images. *Biomed Opt Express.* 2019;10: 5291–5324.

73. Liu X, Cao J, Fu T, et al. Semi-supervised automatic segmentation of layer and fluid region in retinal optical coherence tomography images using adversarial learning. *IEEE Access*. 2019;7:3046–3061.
74. Kugelman J, Alonso–Caneiro D, Read SA, Vincent SJ, Collins MJ. OCT chorio-retinal segmentation with adversarial loss. *2021 Digital Image Computing: Techniques and Applications (DICTA)*. 2021:101–08.
75. Das V, Dandapat S, Bora PK. A data-efficient approach for automated classification of OCT images using generative adversarial network. *IEEE Sensors Letters*. 2020;4:1–4.
76. Lee H, Kim S, Kim MA, Chung H, Kim HC. Post-treatment prediction of optical coherence tomography using a conditional generative adversarial network in age-related macular degeneration. *Retina*. 2021:41.
77. Zhao H, Li H, Maurer-Stroh S, Cheng L. Synthesizing retinal and neuronal images with generative adversarial nets. *Med Image Anal*. 2018;49:14–26.
78. Diaz-Pinto A, Colomer A, Naranjo V, Morales S, Xu Y, Frangi AF. Retinal image synthesis and semi-supervised learning for glaucoma assessment. *IEEE Trans Med Imaging*. 2019;38:2211–2218.
79. Costa P, Galdran A, Meyer MI, et al. End-to-end adversarial retinal image synthesis. *IEEE Trans Med Imaging*. 2018;37:781–791.
80. Shankaranarayana S.M., Ram K., Mitra K., Sivaprakasam M. Joint optic disc and cup segmentation using fully convolutional and adversarial networks. *Fetal, Infant and Ophthalmic Medical Image Analysis*: Springer; 2017:168–176.
81. Lahiri A, Jain V, Mondal A, Biswas PK. Retinal vessel segmentation under extreme low annotation: a Gan based semi-supervised approach. *2020 IEEE International Conference on Image Processing (ICIP)*. 2020:418–422.
82. Sengupta S., Wong A., Singh A., Zelek J., Lakshminarayanan V. DeSupGAN: multi-scale feature averaging generative adversarial network for simultaneous de-blurring and super-resolution of retinal fundus images. In: Fu H, Garvin MK, MacGillivray T, Xu Y, Zheng Y, eds. *Ophthalmic Medical Image Analysis*. 2020:32–41.
83. Andreini P., Bonechi S., Bianchini M., Mecocci A., Scarselli F., Sodi A. A two stage Gan for high resolution retinal image generation and segmentation. *arXiv:1907.12296 [eess.IV]*. 2019.
84. Li PL, O'Neil C, Saberi S, et al. Deep learning algorithm for generating optical coherence tomography angiography (OCTA) maps of the retinal vasculature. *SPIE Optic Eng Appl*. 2020:11511.