



Translating color fundus photography to indocyanine green angiography using deep-learning for age-related macular degeneration screening



Ruoyu Chen^{1,2,6}, Weiyi Zhang^{1,2,6}, Fan Song^{1,2}, Honghua Yu³, Dan Cao³, Yingfeng Zheng⁴✉, Mingguang He^{1,2,5}✉ & Danli Shi^{1,2}✉

Age-related macular degeneration (AMD) is the leading cause of central vision impairment among the elderly. Effective and accurate AMD screening tools are urgently needed. Indocyanine green angiography (ICGA) is a well-established technique for detecting chorioretinal diseases, but its invasive nature and potential risks impede its routine clinical application. Here, we innovatively developed a deep-learning model capable of generating realistic ICGA images from color fundus photography (CF) using generative adversarial networks (GANs) and evaluated its performance in AMD classification. The model was developed with 99,002 CF-ICGA pairs from a tertiary center. The quality of the generated ICGA images underwent objective evaluation using mean absolute error (MAE), peak signal-to-noise ratio (PSNR), structural similarity measures (SSIM), etc., and subjective evaluation by two experienced ophthalmologists. The model generated realistic early, mid and late-phase ICGA images, with SSIM spanned from 0.57 to 0.65. The subjective quality scores ranged from 1.46 to 2.74 on the five-point scale (1 refers to the real ICGA image quality, Kappa 0.79–0.84). Moreover, we assessed the application of translated ICGA images in AMD screening on an external dataset ($n = 13887$) by calculating area under the ROC curve (AUC) in classifying AMD. Combining generated ICGA with real CF images improved the accuracy of AMD classification with AUC increased from 0.93 to 0.97 ($P < 0.001$). These results suggested that CF-to-ICGA translation can serve as a cross-modal data augmentation method to address the data hunger often encountered in deep-learning research, and as a promising add-on for population-based AMD screening. Real-world validation is warranted before clinical usage.

Age-related macular degeneration (AMD) is the leading cause of central vision loss in the aging population¹, mainly consisting of atrophic (“dry”) AMD and neovascular (“wet”) AMD. Dry AMD may progress to wet AMD, which is characterized by central choroidal neovascularization (CNV), resulting in hemorrhaging within the macular region and profound visual

impairment. Effective and accurate AMD screening tools are urgently needed, especially as the aging population intensifies². Over the past few years, applying color fundus (CF) photography to develop deep learning algorithms for automated AMD screening is of great interest^{3–5}. However, it is worth mentioning that only CF images provide limited information in the

¹Experimental Ophthalmology, School of Optometry, The Hong Kong Polytechnic University, Kowloon Hong Kong SAR, China. ²Research Centre for SHARP Vision, The Hong Kong Polytechnic University, Kowloon Hong Kong SAR, China. ³Department of Ophthalmology, Guangdong Academy of Medical Sciences, Guangdong Provincial People’s Hospital, Southern Medical University, Guangzhou, China. ⁴State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, Guangdong Provincial Clinical Research Center for Ocular Diseases, Guangzhou, China. ⁵Centre for Eye and Vision Research (CEVR), 17W Hong Kong Science Park, Hong Kong SAR, China. ⁶These authors contributed equally: Ruoyu Chen, Weiyi Zhang. ✉e-mail: zhyfeng@mail.sysu.edu.hk; mingguang.he@polyu.edu.hk; danli.shi@polyu.edu.hk

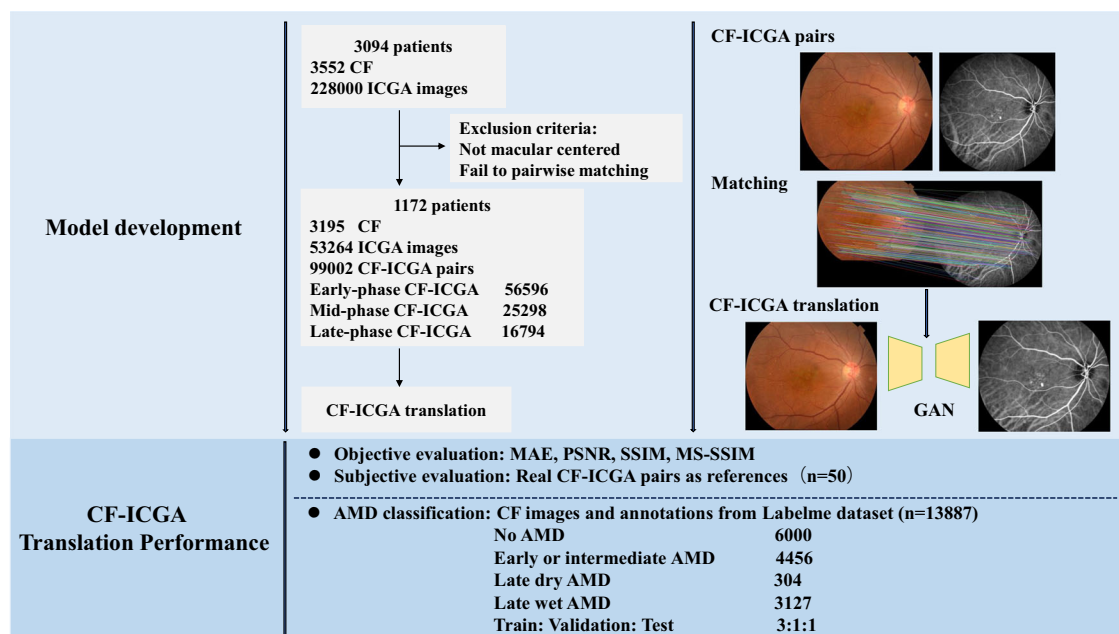


Fig. 1 | Flow chart of the study. GAN generative adversarial networks, CF color fundus photography, ICGA indocyanine green angiography, AMD age-related macular degeneration, MAE mean absolute error, PSNR peak signal-to-noise ratio,

SSIM structural similarity measures, MS-SSIM multi-scale structural similarity measures.

real clinical scenario, because of the unstable image quality and common characteristics shared by several chorioretinal diseases on CF images.

Indocyanine green angiography (ICGA) is a well-established fundus imaging technique for detecting and distinguishing AMD from other chorioretinal diseases⁶⁻⁹. Compared to CF images, ICGA owns its unique advantages in dynamically visualizing deeper choroidal vasculature and lesions behind retinal pigment epithelium^{10,11}. However, ICGA is an invasive imaging modality with potential adverse reactions, including nausea, vomiting, hypotensive shock, etc¹²⁻¹⁵. In addition, the complex operating procedures impede its widespread implementation in clinical settings. Limited and imbalanced ICGA images pose challenges to the development of relevant automated AMD detection models.

Recently, generative adversarial networks (GANs) have showcased remarkable performance in image-to-image translation through two competing types of deep neural networks¹⁶⁻¹⁸, which inspire generating multiple ICGA images from non-invasive CF images. Several exploratory studies have verified the feasibility of cross-modality image translation in the ophthalmic field via GANs, such as CF to fundus fluorescein angiography (FFA) translation^{19,20}. However, few studies have considered detailed phase features and diverse pathological lesions during the process of generating angiographic images, which was essential in diagnosing chorioretinal vasculopathy. Besides, there is no study aimed at achieving CF-to-ICGA translation.

The purpose of our study is to train a GAN-based model for generating realistic ICGA images from CF images using a large clinical dataset and to validate its robustness in AMD screening via external dataset. Our CF-to-ICGA algorithm is expected to provide an effective alternative for addressing the inadequacy and imbalance of ICGA images in deep-learning research, while also promoting the advancement of more accurate AMD screening models.

Results

Images that were not centered on the macula were excluded. The final dataset contains 3195 CF images and 53,264 ICGA images from 1172 patients. Thus, an average of 3 CF images and 45 ICGA images were obtained from each patient. To fully utilize the large number size of ICGA images, we matched each ICGA image with each CF pairwise (paired images were all from the same patient and visit). After excluding images that

failed to be pairwise matched, we finally yielded 99,002 CF-ICGA pairs for model development, in which there were 56596 pairs in early-phase, 25,298 pairs in mid-phase, and 16,794 in late-phase (Fig. 1). The median (interquartile range) age of the participants was 53.04 (± 17.31) years, and 676 (57.7%) were male. The majority of these participants were diagnosed with chorioretinal diseases, including AMD, choroidal neovascularization (CNV), PCV, and pathologic myopia, etc. In the AMD dataset, CF images were classified as no AMD, early or intermediate dry AMD, late dry AMD, and late wet AMD. The study flow chart is shown in Fig. 1. Detailed characteristics of the dataset are presented in Table 1.

Objective evaluation

Pixel-wise comparison between the real and CF-translated ICGA was conducted on the internal test set. For the generated early-phase ICGA, the mean absolute error (MAE), peak signal to noise ratio (PSNR), structural similarity measures (SSIM), and multi-scale structural similarity measures (MS-SSIM) were 86.81, 20.01, 0.57, and 0.68. For the generated mid-phase ICGA, the MAE, PSNR, SSIM, and MS-SSIM were 116.94, 21.74, 0.65, and 0.70. For the generated late-phase ICGA, the MAE, PSNR, SSIM, and MS-SSIM were 118.19, 22.83, 0.57, and 0.74. These results are shown in Table 2.

Subjective evaluation

Example-generated images in the internal and external test sets are shown in Fig. 2. The model efficiently and anatomically achieved CF-ICGA pairwise matching via retinal vascular features, generating realistic ICGA images with detailed structures and lesions. Notably, background noise sometimes present in the real ICGA was learned as unrelated and was effectively excluded in the translated images. The synthesized output images are visually very close to real ones. For the internal test set, a blinded evaluation was conducted using 30 ICGA images with 50% real ICGA images by removing the tags of “Original” and “Generated” on these images. Two experienced ophthalmologists (R.C. and F.S.) selected ICGA images in random order for identification. Among these unlabeled ICGA images, 6.6% and 13.3% generated ICGA images could be differentiated from the real ones. The distinguishing features encompassed the blurry boundary of lesions, strange lesions that are apparently against established clinical knowledge, vascular discontinuity, and the blurry texture of choroidal

Table 1 | Dataset characteristics

	<i>N</i>	<i>No</i>					
Model development		normal	AMD	PCV	CNV	PM	Other
Patients	1172	267 (22.8%)	320 (27.3%)	158 (13.5%)	141 (12.0%)	104 (8.9%)	182 (15.5%)
External validation							
CF-ICGA pairs	50	Dry AMD	Wet AMD	PCV	CSC	Others	
		12 (24.0%)	15 (30.0%)	5 (10.0%)	6 (12.0%)	12 (24.0%)	
AMD Classification (AMD Dataset)		No AMD	Early or intermediate dry AMD	Late dry AMD	Late wet AMD		
Train	8331	3600 (43.2%)	2673 (32.1%)	182 (2.2%)	1876 (22.5%)		
Validation	2781	1200 (43.2%)	891 (32.0%)	61 (2.2%)	629 (22.6%)		
Test	2775	1200 (43.2%)	892 (32.1%)	61 (2.2%)	622 (22.4%)		
Total	13887	6000 (43.2%)	4456 (32.1%)	304 (2.2%)	3127 (22.5%)		

CF color fundus photography, ICGA indocyanine green angiography, AMD age-related macular degeneration, PCV polypoidal choroidal vasculopathy, CNV choroidal neovascularization, CSC central serous chorioretinopathy, *N* number of patients/images.

Table 2 | Objective evaluation of real and translated indocyanine green angiography (ICGA) images

	MAE	PSNR	SSIM	MS-SSIM
Early-phase ICGA	86.81	20.01	0.57	0.68
Mid-phase ICGA	116.94	21.74	0.65	0.70
Late-phase ICGA	118.19	22.83	0.57	0.74

MAE mean absolute error, PSNR peak signal-to-noise ratio, SSIM structural similarity measures, MS-SSIM multi-scale structural similarity measures.

vessels (Supplementary Fig. 5). For the external test set, the dissimilarities in image characteristics between the training dataset and the external dataset were readily apparent. Consequently, the real and generated ICGA images could be discerned based on image style alone, and nearly all the generated ICGA images could be distinguished from the authentic ones (as illustrated in Fig. 2, 4th row). Thus, we did not conduct blind evaluations in this case.

Image quality assessment, which considers chorioretinal structure and lesions, was based on a five-point scale. The mean (standard deviation [SD]) of the scores for early-phase ICGA was 1.46 (0.76), 2.08 (0.85) for the internal and external test set respectively, assessed by the first grader, and 1.48 (0.76), 2.12 (0.77) by the second grader. The mean (SD) of the scores for mid-phase ICGA were 2.02 (0.84) and 2.56 (0.91) in the internal and external test sets respectively, assessed by the first grader, and 1.94 (0.89), 2.66 (0.98) by the second grader. The mean (SD) of the scores for late-phase ICGA were 2.04 (0.81) and 2.74 (0.82) in the internal and external test sets respectively, assessed by the first grader, and 1.96 (0.78), 2.64 (0.74) by the second grader. Cohen’s kappa values indicate an excellent agreement between the two graders for assessing image quality, with the Kappa value of 0.79, 0.81 for early-phase ICGA, 0.81, 0.80 for mid-phase ICGA, and 0.81, 0.84 for late-phase ICGA in the internal and external test sets, respectively (Table 3). This reflects the high quality of synthesized images for anatomical features (vessel, optic disc, and macula) and lesions (drusen, choroidal neovascular, atrophy, subretinal fluid, and hemorrhage).

In the internal and external test sets, 5.3% and 12.0% of the generated ICGA images were of poor quality (>=4 points) due to the following reasons: blurry CF images impact the quality of synthetic ICGA. Besides, choroidal lesions could be sheltered by extremely thick subretinal hemorrhage and scar, resulting in the generation of false negative or positive lesions, as demonstrated in Supplementary Fig. 3.

While this study primarily focused on AMD, Supplementary Fig. 4 provides generation examples for normal fundus and other diseases, such as

polypoidal choroidal vasculopathy, central serous chorioretinopathy, pathologic myopia, and punctate inner choroidopathy to further demonstrate the comprehensiveness of this model.

AMD classification

The addition of generated ICGA on top of CF significantly improved the accuracy of AMD classification on the inhouse AMD datasets²¹, as illustrated in Table 4 and Fig. 3. The integration of generated ICGA images significantly reduced error rates for AMD categories (Fig. 4). In the test set, 219 (18.3%) non-AMD cases, 180 (20.2%) early or intermediate dry AMD cases, 5 (8.2%) late dry AMD cases, and 49 (7.9%) late wet AMD cases were falsely predicted based on CF alone. Among these false prediction cases, 21.0%–48.9% with no AMD, 52.2–61.1% with early or intermediate dry AMD, 20.0–60.0% with late dry AMD, and 42.9–44.9% with late wet AMD could be accurately predicted after adding generated ICGA images (Table 5 and Supplementary Fig. 6). In addition, significant differences in AMD classification accuracies were found between CF alone and CF+generated early-phase ICGA+generated mid-phase ICGA, as well as CF alone and CF+generated early-phase ICGA+generated mid-phase+generated late-phase ICGA ($P < 0.001$). However, there were no statistical differences between CF alone and CF+early-phase ICGA ($P = 0.44$) (Table 4). In general, these results demonstrated that incorporating synthetic ICGA images with CF significantly enhances AMD classification accuracy.

Discussion

In the present study, we innovatively demonstrated that high-resolution ICGA images could be synthesized based on CF images using GANs. The reliability of this model in generating authentic ICGA images has been proven through both internal and external validation. Additionally, we have illustrated that the integration of the translated ICGA images with CF images significantly improves the accuracy of AMD screening. Our study not only established the feasibility of predicting choroidal abnormalities more accurately from the more accessible CF images via GANs but also introduced a cross-modality approach to augment data for AMD-related deep learning research.

Image-to-image translation has garnered significant attention within the domain of fundus multimodal imaging systems^{16,22,23}. The generation of realistic ICGA images from CF images posed a significant challenge due to the masking effect of the retinal pigment epithelium and the intricate anatomical structure of the choroid. To address this challenge, our algorithm leveraged the power of Pix2pix HD, which belongs to the conditional GANs family, and demonstrates excellent performance in image-to-image translation via reducing adversarial loss and pixel-reconstruction error²⁴. Its strong ability in denoising, super-resolution, and feature extraction has

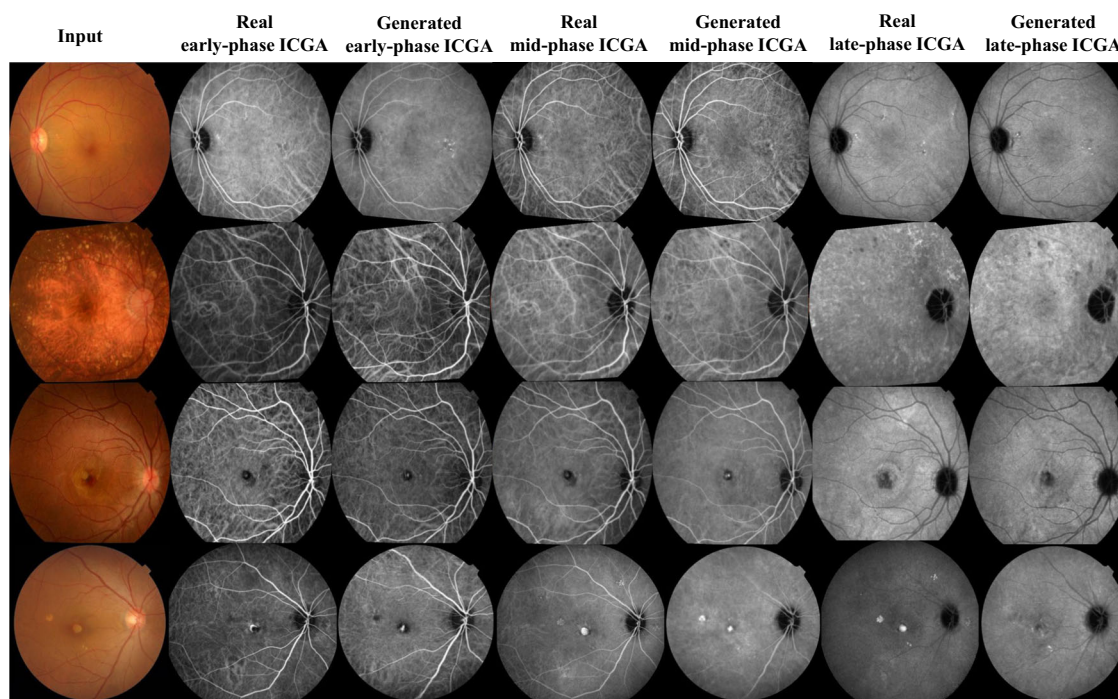


Fig. 2 | Examples of real and translated indocyanine green angiography (ICGA). 1st row, early dry age-related macular degeneration (AMD), 2nd row, intermediate dry AMD, 3rd row, wet AMD, 4th row, wet AMD. 1 – 3 rows: internal test set, CF

were registered with ICGA, rotation occurs during this process. 4th row: external test set.

Table 3 | Subjective evaluation of real and translated indocyanine green angiography (ICGA) image quality

	Internal test set (N = 50)			External test set (N = 50)		
	Rater 1 Mean (SD)	Rater 2 Mean (SD)	Kappa	Rater 1 Mean (SD)	Rater 2 Mean (SD)	Kappa
early-phase ICGA	1.46 (0.76)	1.48 (0.76)	0.79	2.08 (0.85)	2.12 (0.77)	0.81
mid-phase ICGA	2.02 (0.84)	1.94 (0.89)	0.81	2.56 (0.91)	2.66 (0.98)	0.80
late-phase ICGA	2.04 (0.81)	1.96 (0.78)	0.81	2.74 (0.82)	2.64 (0.74)	0.84

SD standard deviation.

Table 4 | Age-related macular degeneration (AMD) classification based on color fundus photography (CF) and CF+ translated indocyanine green angiography (ICGA) images on the AMD dataset (n = 13887)

	F1-score	Sensitivity	Specificity	Accuracy	AUC	P value*
CF	0.8386	0.8368	0.9323	0.8368	0.9312	
CF+early	0.8601	0.8598	0.9428	0.8598	0.9407	0.4400
CF+early+mid	0.8854	0.8850	0.9466	0.8850	0.9632	<0.0001**
CF+early+mid+late	0.8875	0.8872	0.9474	0.8872	0.9688	<0.0001**

*The distinctions of ROC curves were evaluated using the bootstrap method in the R package pROC between the following scenarios: CF alone versus CF+generated early-phase ICGA images, CF alone versus CF+generated early-phase ICGA+generated mid-phase ICGA, CF alone versus generated early-phase ICGA+generated mid-phase ICGA+generated late-phase ICGA. **P < 0.05. The AMD dataset, train: validation: test = 6:2:2.

provided a robust foundation for accurate CF-to-ICGA translation²⁵. Considering exudation, hemorrhage, and other abnormalities may result in poor generation, we applied an additional Gradient Variance Loss to generate high-resolution details with sharp edges by minimizing the distance between the computed variance maps and enforcing the model to produce high-variance gradient maps²⁶.

ICGA is a unique modality for detecting choroidal abnormalities because of its strong penetration and high contrast characteristics, offering

much more information on choroidal circulation than other non-invasive approaches, such as CF and optic coherence tomography angiography (OCTA)^{27–29}. In the current study, the detailed phase information of ICGA images, such as cut-off time and fluorescence characteristics of each phase were considered in model training. The current GAN-based model could authentically generate ICGA images of early, medium, and late phases respectively from a single CF image, suggesting an effective alternative for observing choroidal lesions dynamically and non-invasively. Our results

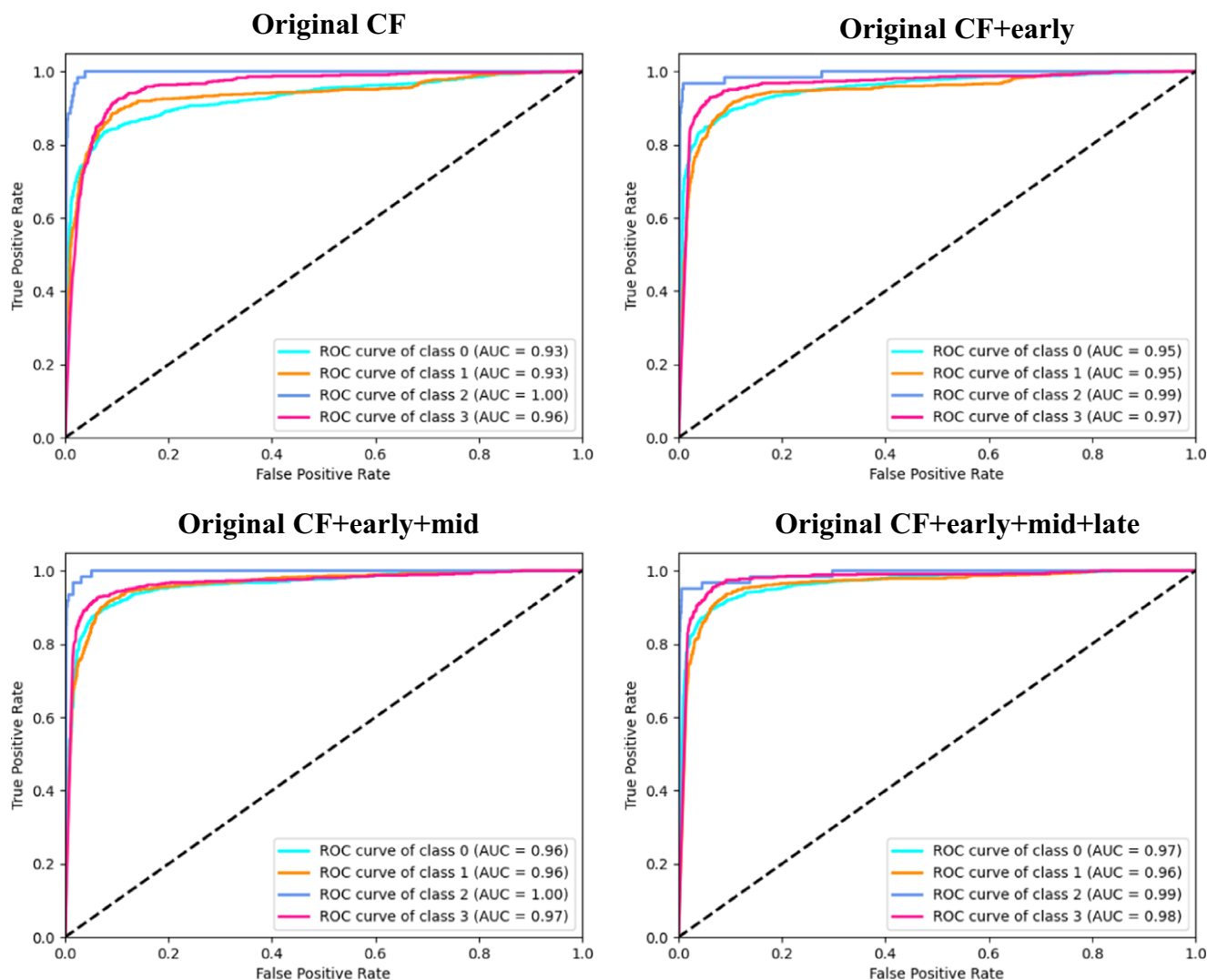


Fig. 3 | ROC curves of age-related macular degeneration (AMD) classification on the AMD dataset ($n = 13887$) with and without the addition of translated indocyanine green angiography (ICGA) on top of color fundus photography (CF). The classification is based on four categories: 0 = no AMD, 1 = early or intermediate dry AMD, 2 = late dry AMD, and 3 = late wet AMD. 1st row

(left) = Original CF, 1st row (right) = Original CF plus generated early-phase ICGA images, 2nd row (left) = Original CF plus generated early-phase plus mid-phase ICGA images, 2nd row (right) = Original CF plus generated early-phase plus mid-phase plus late-phase ICGA images.

illustrated that the addition of generated ICGA images significantly improves the accuracy of AMD classification within an external dataset compared to CF alone. The comparable performance between CF and CF plus early-phase ICGA may result from the limited information provided by early-phase ICGA alone.

AMD shares many common features with PCV, but significant disparities in treatment response and prognosis are also found in these two diseases, emphasizing the need for distinguishing PCV from AMD^{30,31}. In our study, the branch vascular networks and polyps, which are the landmarks for diagnosing PCV, can be realistically generated (Supplementary Fig. 4, 1st row). In addition, normal fundus and characteristic lesions of other chorioretinal diseases could also be well translated on ICGA images (Supplementary Fig. 4, 2rd–5th row). Our findings also demonstrated that 42.9% to 60.0% of the cases misclassified by CF alone could be correctly predicted after adding all-phase generated ICGA, further showcasing that these synthesized ICGA can effectively reduce classification errors and facilitate more accurate AMD screening in the current research background. Nevertheless, generative technology is still at the exploratory stage in ophthalmic research with yet unclear clinical relevance^{16,32}. Prospective trials and clinician-engineer

collaboration are necessary to prove whether the application of synthetic images will optimize clinical practice.

The inadequacy of ICGA images represents a challenge in deep learning research aimed at developing automated tools for detecting chorioretinal conditions. In the past few decades, GANs-based technology has been introduced to expand datasets and protect patient privacy via generating realistic images, especially of uncommon diseases^{33–35}. Several studies have reported that integrating synthetic images could enhance the performance of machine-learning models in the ophthalmic field^{36–38}. In addition, synthetic data could be incorporated into various research tasks, such as lesion segmentation, image denoising, super-resolution, etc¹⁶. Thus, translated ICGA images may also be expected to address data shortages and imbalances, enabling low-shot or zero-shot training in ICGA relevant deep-learning models. The cross-modality image translation may be a good choice in maximizing the potential of existing datasets for the development of deep learning systems.

This research also has limitations. Firstly, the presence of extremely thick subretinal hemorrhage and scars, along with blurry CF images can potentially impact the quality of the generated ICGA images, resulting in false negative or false positive lesions due to incorrect matching of blurry

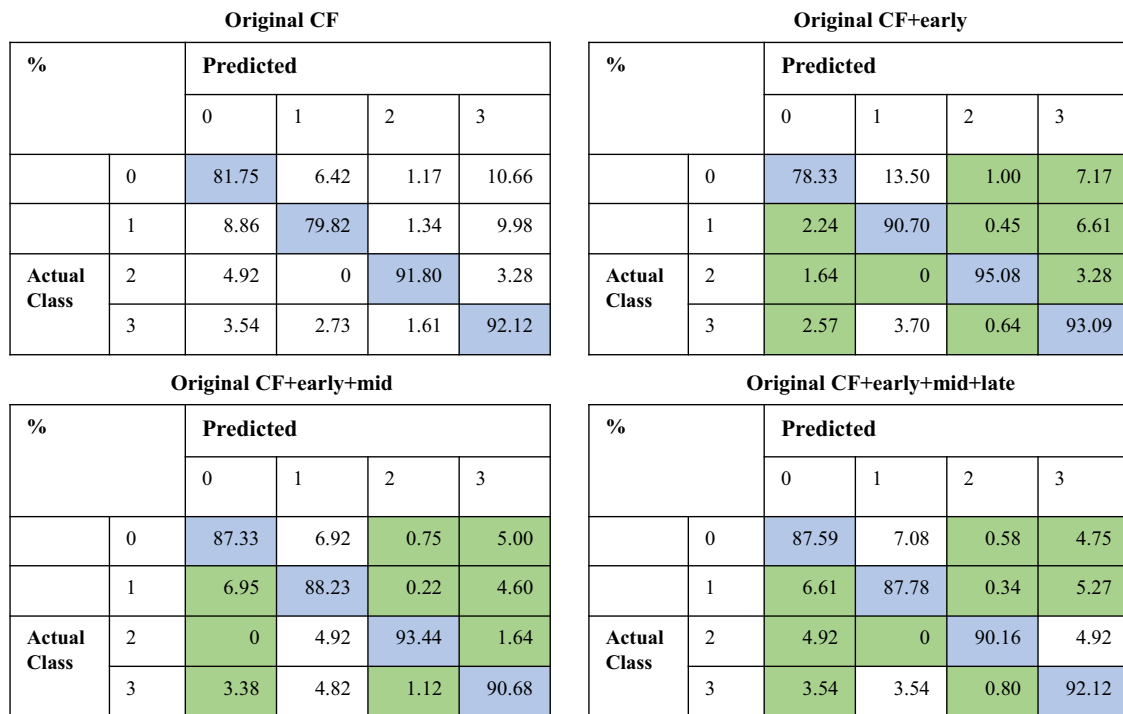


Fig. 4 | Confusion matrix of age-related macular degeneration (AMD) classification results on the AMD dataset ($n = 13887$) with and without the addition of translated indocyanine green angiography (ICGA) on top of color fundus photography (CF). The classification is based on four categories: 0 = no AMD, 1 = early or intermediate dry AMD, 2 = late dry AMD, and 3 = late wet AMD. 1st row

(left) = Original CF, 1st row (right) = Original CF plus generated early-phase ICGA images, 2nd row (left) = Original CF plus generated early-phase plus mid-phase ICGA images, 2nd row (right) = Original CF plus generated early-phase plus mid-phase plus late-phase ICGA images.

images in translation procedure. Secondly, though a large-scale dataset was utilized for training in the current study, more variable real-world datasets are critical for improving the diversity and applicability of generated ICGA images. Most importantly, the current research aims to explore the authenticity and limitations of synthetic images in a controlled research context, future validation is warranted to evaluate whether synthetic ICGA images could assist clinical decision-making.

In conclusion, we innovatively developed a deep-learning model for generating ICGA images of early, medium, and late phases from CF images. The algorithm showed high authenticity in generating

anatomical structures and pathological lesions in both internal and external datasets. These findings highlight the potential of the CF-to-ICGA model as a valuable approach for evaluating chorioretinal circulation and abnormalities non-invasively, as well as a promising tool for overcoming data shortages in machine-learning model training. Further clinical trials are required to translate this research discovery into clinical benefit in real-world practice.

Methods

Data

We collected a total of 3552 CF images and 228,000 ICGA images from a tertiary center between 2016 and 2019. These images were sourced from 3094 patients who had been diagnosed with a range of ocular diseases, including age-related macular degeneration (AMD), polypoidal choroidal vasculopathy (PCV), choroidal neovascularization (CNV), central serous chorioretinopathy (CSCR), pathologic myopia (PM), and ocular inflammatory diseases, etc. All patient data underwent anonymization and de-identification processes. The CF images were captured using Topcon TRC-50XF and Zeiss FF450 Plus (Carl Zeiss, Inc., Jena, Germany) cameras, with resolutions spanning from 1110×1467 to 2600×3200 . The ICGA images were acquired using Heidelberg Spectralis cameras (Heidelberg, Germany) at a resolution of 768×768 .

To assess the reliability of our model, we conducted a retrospective collection of 50 paired CF and ICGA images specifically focusing on choroidal conditions from Guangdong Provincial People's Hospital. The CF images were captured using Topcon TRC-50XF cameras, while the ICGA images were obtained using Heidelberg Spectralis cameras from Heidelberg, Germany.

To assess the efficacy of translated ICGA images in enhancing AMD detection, we utilized a separated dataset procured from a web-based, cloud-sourcing platform situated in Guangzhou, China²¹. This dataset encompassed macular and disc-centered CF images sourced from 36

Table 5 | Comparison of false prediction cases on different age-related macular degeneration (AMD) categories based on color fundus photography (CF) only but were correctly predicted after the addition of translated indocyanine green angiography (ICGA) images in the test set of AMD dataset ($n = 2775$)

	No AMD	Early or intermediate dry AMD	Late dry AMD	Late wet AMD
Total, N	1200	892	61	622
False prediction, N (%)				
CF only	219 (18.3)	180 (20.2)	5 (8.2)	49 (7.9)
Correct prediction rate after adding generated ICGA images, N (%)				
CF+early	46 (21.0)	110 (61.1)	2 (40.0)	22 (44.9)
CF+early+mid	107 (48.9)	94 (52.2)	1 (20.0)	21 (42.9)
CF+early+mid+late	101 (46.1)	94 (52.2)	3 (60.0)	21 (42.9)

N number of cases, CF+early CF+generated early-phase ICGA images, CF+mid CF+generated mid-phase ICGA images, CF+late CF+generated late-phase ICGA images.

ophthalmology departments, optometry clinics, and screening facilities spanning various regions of China. The grading of AMD followed the Beckman clinical classification system which involved categorizing patients into groups based on the absence of AMD, early-stage AMD, intermediate-stage AMD, and late-stage AMD (atrophic or neovascular), considering the severity and pathological progression³⁹. A detailed overview of the dataset characteristics is presented in Table 1.

The study adheres to the tenets of the Declaration of Helsinki. The Institutional Review Board approved the study and individual consent for this retrospective analysis was waived (Approval Number: 2021KYPJ164 – 3).

CF and ICGA matching

We conducted a pairwise matching process between CF and ICGA images obtained from the same eye and visit. CF and ICGA images captured from the same eye and eye position share common retinal vessel features, which are stable and easily detectable in both modalities. Thus, we utilized the retinal vessel features to achieve CF-ICGA pairwise matching. We applied the Retina-based Microvascular Health Assessment System (RMHAS) segmentation module to extract retinal vessels from CF and ICGA images⁴⁰. The cross-modality matching process follows previous work⁴¹. Initially, key points from both modalities were identified from the corresponding vessel maps using the AKAZE key point detector through feature matching⁴². Then the RANSAC (random sample consensus) algorithm was utilized to generate homography matrices and eliminate outliers to facilitate registration⁴³. In addition, we implemented a validity restriction to ensure the accuracy of the registration. This restriction enforced a rotation scale value between 0.8 and 1.3, and an absolute rotation radian value less than 2 before the warping transformation. Additionally, we filtered out image pairs with poor registration performance by setting a dice coefficient threshold of 0.5, which was determined empirically based on our dataset during experiment.

CF to ICGA translation

In our training process, we utilized CF images as input and the corresponding early-phase, mid-phase, and late-phase ICGA images as the ground truth to train three separate models with CF-ICGA pairs. To reduce variation, we established specific time ranges for each phase of ICGA images: 25 s to 3 min for early-phase, 3 to 15 min for mid-phase, and after 15 min for late-phase. The dataset was split into three sets: training, validation, and testing, following a 6:2:2 ratio based on patient level. During training, the images were resized to 512 × 512 and fed into pix2pixHD⁴⁴, a popular GAN model consisting of generator G, which synthesizes candidate samples based on the data distribution of the original dataset, and a discriminator D, which distinguishes the synthesized candidate samples from the original samples. The discriminator employed a multi-scale convolutional neural network that divided the image into patches and evaluated the fidelity of each patch. This approach contributed to the generation of high-resolution ICGA images that closely resembled the real ones. To enhance the generation of high-frequency components, such as chorioretinal structure and lesions, we incorporated Gradient Variance Loss to prevent overfitting²⁶. Additionally, extensive data augmentations were applied during training, including random resized crops within a scale range of 0.3–3.5, random horizontal or vertical flipping, and random rotations within a range of 0–45 degrees. Individual CF images and their corresponding generated ICGA images all underwent the same augmentation process. The models were trained with a batch size of 4 and a learning rate of 0.0002. Each training session was preset to run for a total of 50 epochs, ensuring an adequate number of iterations for convergence and optimization in our experiments.

Assessment of CF-to-ICGA Translation Performance.

Objective evaluation

For the evaluation of image authenticity, we employed four standard objective measures widely used in image generation for our internal test set. MAE computes the average absolute pixel difference between the

generated image and the corresponding real image. It quantifies the overall discrepancy in pixel values, indicating the level of fidelity in generating accurate details. PSNR is the approximation of human perception regarding reconstruction quality. It measures the ratio between the maximum possible power of a signal and the power of the noise interfering with it. SSIM⁴⁵ assesses the structural similarity between images, with a value of 1 representing complete similarity and 0 indicating no similarity. SSIM provides insights into the visual resemblance and coherence between the generated and real images. MS-SSIM⁴⁶ supplies more flexibility in incorporating the variations of viewing conditions and image resolution. The higher the SSIM, MS-SSIM, and PSNR, the better the quality of the generated images.

Subjective evaluation

Fifty images from the internal and external test sets were randomly assigned to two experienced ophthalmologists (R.C. and F.S.) for visual quality assessment. The ophthalmologists assessed the translated images subjectively, considering factors such as the global similarity, the fidelity of anatomical structures, and the depiction of fluorescence-based pathological lesions, on a scale of 1 to 5 (1 = excellent, 2 = good, 3 = normal, 4 = poor, and 5 = very poor), with score 1 referring to the image quality of the real ICGA image. The detailed grading criteria and examples of different quality were demonstrated in Supplementary Figs. 1 and 2. To determine the agreement between the ophthalmologists, we calculated Cohen's linearly weighted kappa score⁴⁷. This score ranges from –1 to 1, with values between 0.40 and 0.60 indicating moderate agreement, 0.60 and 0.80 indicating substantial agreement, and 0.80 and 1.00 indicating almost perfect agreement. The inter-rater agreement was assessed based on this evaluation metric. Besides, blind assessment was conducted using ICGA images without tags of “Original” and “Generated” to investigate whether it is possible to reliably distinguish between real and generated ICGA images for ophthalmologists.

AMD classification

CF-ICGA translation performance was also evaluated using an external AMD dataset, which consists of real CF images and AMD annotations. Our GAN-based model generated corresponding early-phase, mid-phase, and late-phase ICGA images respectively from each real CF image. We did experiments using the Swin-transformer to explore whether the addition of ICGA images generated by our model could improve the classification accuracy of AMD⁴⁸. The same hyperparameters were set in each experiment to classify AMD based on CF, CF + generated early-phase ICGA, CF + generated early-phase + mid-phase ICGA images, and CF + generated early-phase + mid-phase + late-phase ICGA images. Four relevant embedding models were initialized with pretrained weights from ImageNet⁴⁹, and the classifier shared the same training data with the embedding model. Specifically, the features from different images were extracted by Swin-transformer into 512-dimensional embeddings, these embeddings were then concatenated and passed through a fully-connected layer, followed by a softmax layer to obtain the classification output. The dataset was divided into training, validation, and test sets in a ratio of 6:2:2. During training, the images were resized to 512 × 512 and augmented with random horizontal flips and rotations ranging from –30 to 30 degrees. The Adam optimizer with a learning rate of 1e-5 and a batch size of 4 was employed. Each experiment was trained for 30 epochs, and the models with the highest area under the ROC curve (AUC) value on the validation set were selected for testing. The performance of AMD classification was assessed on the test set using various metrics, including F1-score, Sensitivity, Specificity, Accuracy, and AUC. Confusion matrices were utilized to analyze the fine-grained class-level performance as well. Additionally, to verify whether the addition of generated ICGA images could improve AMD prediction performance, we conducted a comparative analysis of ROC curves using the bootstrap method in the R

package pROC⁵⁰. Specifically, we examined the distinctions of ROC curves between the following scenarios: CF alone versus CF+generated early-phase ICGA images, CF alone versus CF+generated early-phase ICGA+generated mid-phase ICGA, CF alone versus generated early-phase ICGA+generated mid-phase ICGA+generated late-phase ICGA. Reported *P* values are two-sided and the results were considered statistically significant with *P* < 0.05.

Data availability

The data used for model development of this study are not openly available due to reasons of privacy and are available from the corresponding author upon reasonable request. The AMD dataset used for external validation is located on a controlled access data platform. Interested researchers can contact M.H. (mingguang.he@polyu.edu.hk) for more information.

Code availability

The deep-learning model was developed using PyTorch (<http://pytorch.org>). We trained the model on an NVIDIA GeForce RTX 3090 card. The code for deep learning model development can be accessed at <https://github.com/NVIDIA/pix2pixHD>.

Received: 22 July 2023; Accepted: 18 January 2024;

Published online: 12 February 2024

References

1. Thomas, C. J., Mirza, R. G. & Gill, M. K. Age-related macular degeneration. *Med. Clin. North Am.* **105**, 473–491 (2021).
2. Li, J. Q. et al. Prevalence and incidence of age-related macular degeneration in Europe: a systematic review and meta-analysis. *Br. J. Ophthalmol.* **104**, 1077–1084 (2020).
3. Burlina, P. M. et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol.* **135**, 1170–1176 (2017).
4. Grassmann, F. et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology* **125**, 1410–1420 (2018).
5. Burlina, P. M. et al. Use of deep learning for detailed severity characterization and estimation of 5-year risk among patients with age-related macular degeneration. *JAMA Ophthalmol.* **136**, 1359–1366 (2018).
6. Stanga, P. E., Lim, J. I. & Hamilton, P. Indocyanine green angiography in chorioretinal diseases: indications and interpretation: an evidence-based update. *Ophthalmology* **110**, 15–21 (2003).
7. Tan, C. S., Ngo, W. K., Chen, J. P., Tan, N. W. & Lim, T. H. EVEREST study report 2: imaging and grading protocol, and baseline characteristics of a randomised controlled trial of polypoidal choroidal vasculopathy. *Br. J. Ophthalmol.* **99**, 624–628 (2015).
8. Eandi, C. M. et al. Indocyanine green angiography and optical coherence tomography angiography of choroidal neovascularization in age-related macular degeneration. *Investig. Ophthalmol. Vis. Sci.* **58**, 3690–3696 (2017).
9. Guyer, D. R. et al. Classification of choroidal neovascularization by digital indocyanine green videoangiography. *Ophthalmology* **103**, 2054–2060 (1996).
10. Landsman, M. L., Kwant, G., Mook, G. A. & Zijlstra, W. G. Light-absorbing properties, stability, and spectral stabilization of indocyanine green. *J. Appl. Physiol.* **40**, 575–583 (1976).
11. Desmettre, T., Devoisselle, J. M. & Mordon, S. Fluorescence properties and metabolic features of indocyanine green (ICG) as related to angiography. *Surv. Ophthalmol.* **45**, 15–27 (2000).
12. Hope-Ross, M. et al. Adverse reactions due to indocyanine green. *Ophthalmology* **101**, 529–533 (1994).
13. Bonte, C. A., Ceuppens, J. & Leys, A. M. Hypotensive shock as a complication of infracyanine green injection. *Retina* **18**, 476–477 (1998).
14. Olsen, T. W., Lim, J. I., Capone, A. Jr., Myles, R. A. & Gilman, J. P. Anaphylactic shock following indocyanine green angiography. *Arch. Ophthalmol.* **114**, 97 (1996).
15. Engel, E. et al. Light-induced decomposition of indocyanine green. *Investig. Ophthalmol. Vis. Sci.* **49**, 1777–1783 (2008).
16. You, A., Kim, J. K., Ryu, I. H. & Yoo, T. K. Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey. *Eye Vis. (Lond., Engl.)* **9**, 6 (2022).
17. Gong, M., Chen, S., Chen, Q., Zeng, Y. & Zhang, Y. Generative adversarial networks in medical image processing. *Curr. Pharm. Des.* **27**, 1856–1868 (2021).
18. Jeong, J. J. et al. Systematic review of generative adversarial networks (GANs) for medical image classification and segmentation. *J. Digit. Imaging* **35**, 137–152 (2022).
19. Li, P. et al. Synthesizing multi-frame high-resolution fluorescein angiography images from retinal fundus images using generative adversarial networks. *Biomed. Eng. online* **22**, 16 (2023).
20. Shi, D. et al. Translation of color fundus photography into fluorescein angiography using deep learning for enhanced diabetic retinopathy screening. *Ophthalmol. Sci.* **3**, 100401 (2023).
21. Keel, S. et al. Development and validation of a deep-learning algorithm for the detection of neovascular age-related macular degeneration from colour fundus photographs. *Clin. Exp. Ophthalmol.* **47**, 1009–1018 (2019).
22. Yu, Z. et al. Retinal image synthesis from multiple-landmarks input with generative adversarial networks. *Biomed. Eng. Online* **18**, 62 (2019).
23. Chen, J. S. et al. Deepfakes in ophthalmology: applications and realism of synthetic retinal images from generative adversarial networks. *Ophthalmol. Sci.* **1**, 100079 (2021).
24. Isola, P., Zhu, J., Zhou, T. & Efros, A. A. Image-to-Image Translation with Conditional Adversarial Networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 5967–5976 (2016).
25. Zhao, J., Hou, X., Pan, M. & Zhang, H. Attention-based generative adversarial network in medical imaging: a narrative review. *Comput. Biol. Med.* **149**, 105948 (2022).
26. Abrahamyan, L., Truong, A.M., Philips, W. & Deligiannis, N. Gradient variance loss for structure-enhanced image super-resolution. in *ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 3219–3223 (2022).
27. Arrigo, A. et al. Morphological and functional relationship between OCTA and FA/ICGA quantitative features in AMD-related macular neovascularization. *Front. Med.* **8**, 758668 (2021).
28. Azar, G. et al. Polypoidal choroidal vasculopathy diagnosis and neovascular activity evaluation using optical coherence tomography angiography. *BioMed. Res. Int.* **2021**, 1637377 (2021).
29. Stahl, A. The diagnosis and treatment of age-related macular degeneration. *Dtsch. Arztebl. Int.* **117**, 513–520 (2020).
30. Wong, C. W. et al. Age-related macular degeneration and polypoidal choroidal vasculopathy in Asians. *Prog. Retin. Eye Res.* **53**, 107–139 (2016).
31. Anantharaman, G. et al. Polypoidal choroidal vasculopathy: pearls in diagnosis and management. *Indian J. Ophthalmol.* **66**, 896–908 (2018).
32. Wang, Z. et al. Generative adversarial networks in ophthalmology: what are these and how can they be used? *Curr. Opin. Ophthalmol.* **32**, 459–467 (2021).
33. Lim, G., Thombre, P., Lee, M. L. & Hsu, W. Generative data augmentation for diabetic retinopathy classification. in *2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)* 1096–1103 (2020).

34. Qasim, A. B. et al. Red-GAN: Attacking class imbalance via conditioned generation. Yet another medical imaging perspective. *ArXiv, abs/2004.10734* (2020).
35. Liu, T. Y. A., Farsiu, S. & Ting, D. S. Generative adversarial networks to predict treatment response for neovascular age-related macular degeneration: interesting, but is it useful?. *Br. J. Ophthalmol.* **104**, 1629–1630 (2020).
36. Zheng, C. et al. Assessment of generative adversarial networks model for synthetic optical coherence tomography images of retinal disorders. *Transl. Vis. Sci. Technol.* **9**, 29 (2020).
37. Abdelmotaal, H., Abdou, A. A., Omar, A. F., El-Sebaity, D. M. & Abdelazeem, K. Pix2pix conditional generative adversarial networks for scheinpflug camera color-coded corneal tomography image generation. *Transl. Vis. Sci. Technol.* **10**, 21 (2021).
38. Yoo, T. K., Choi, J. Y., Kim, H. K., Ryu, I. H. & Kim, J. K. Adopting low-shot deep learning for the detection of conjunctival melanoma using ocular surface images. *Comput. Methods Prog. Biomed.* **205**, 106086 (2021).
39. Ferris, F. L. 3rd et al. Clinical classification of age-related macular degeneration. *Ophthalmology* **120**, 844–851 (2013).
40. Shi, D. et al. A Deep Learning System for Fully Automated Retinal Vessel Measurement in High Throughput Image Analysis. *Front. Cardiovasc. Med.* **9**, 823436 (2022).
41. Shi, D., He, S., Yang, J., Zheng, Y. & He, M. One-shot retinal artery and vein segmentation via cross-modality pretraining. *Ophthalmol. Sci.* **4**, 100363 (2023).
42. Alcantarilla, P. F. Fast explicit diffusion for accelerated features in nonlinear scale spaces. In *British Machine Vision Conference (BMVC)* (2013).
43. Fischler, M. A. & Bolles, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **24**, 381–395 (1981).
44. Wang, T. et al. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 8798–8807 (2017).
45. Zhou, W., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
46. Wang, Z., Simoncelli, E. P. & Bovik, A. C. Multiscale structural similarity for image quality assessment. In *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers* 2, 1398–1402 (2003).
47. Koch, L. G. G. The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174 (1977).
48. Liu, Z. et al. Swin transformer: hierarchical vision transformer using shifted windows. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 9992–10002 (2021).
49. Russakovsky, O. et al. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2014).
50. Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* **12**, 77 (2011).

Acknowledgements

Danli Shi and Mingguang He disclose support for the research and publication of this work from the Start-up Fund for RAPs under the Strategic Hiring Scheme (Grant Number: P0048623) and the Global STEM Professorship Scheme (Grant Number: P0046113) from HKSAR. Yingfeng Zheng discloses support from the National Natural Science Foundation of China (Grant Number: 82171034). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

D.S. and M.H. conceived the study. D.S. and W.Z. built the deep learning model. D.S. and R.C. did the literature search and analysed the data. M.H., Y.Z., H.Y. and D.C. provided data. D.S., R.C., Y.Z., F.S. and M.H. contributed to key data interpretation. R.C. is a major contributor in writing the manuscript. All authors critically read and revised the manuscript. R.C. and W.Z. contributed equally in this study. D.S., M.H. and Y.Z. are corresponding authors and contributed equally.

Competing interests

M.H. and D.S. are inventors of the technology mentioned in the study patented as “A method to translate fundus photography to realistic angiography” (CN115272255A). All other authors declare no financial or non-financial competing interests.

Additional information

Supplementary information The online version contains

supplementary material available at <https://doi.org/10.1038/s41746-024-01018-7>.

Correspondence and requests for materials should be addressed to Yingfeng Zheng, Mingguang He or Danli Shi.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024