

ARTICLE OPEN



Ultra-wide field and new wide field composite retinal image registration with AI-enabled pipeline and 3D distortion correction algorithm

Fritz Gerald P. Kalaw^{1,2,3}, Melina Cavichini^{1,2}, Junkang Zhang⁴, Bo Wen⁴, Andrew C. Lin², Anna Heinke^{1,2}, Truong Nguyen⁴, Cheolhong An⁴, Dirk-Uwe G. Bartsch¹, Lingyun Cheng¹ and William R. Freeman^{1,2,3,4}✉

© The Author(s) 2023

PURPOSE: This study aimed to compare a new Artificial Intelligence (AI) method to conventional mathematical warping in accurately overlaying peripheral retinal vessels from two different imaging devices: confocal scanning laser ophthalmoscope (cSLO) wide-field images and SLO ultra-wide field images.

METHODS: Images were captured using the Heidelberg Spectralis 55-degree field-of-view and Optos ultra-wide field. The conventional mathematical warping was performed using Random Sample Consensus—Sample and Consensus sets (RANSAC-SC). This was compared to an AI alignment algorithm based on a one-way forward registration procedure consisting of full Convolutional Neural Networks (CNNs) with Outlier Rejection (OR CNN), as well as an iterative 3D camera pose optimization process (OR CNN + Distortion Correction [DC]). Images were provided in a checkerboard pattern, and peripheral vessels were graded in four quadrants based on alignment to the adjacent box.

RESULTS: A total of 660 boxes were analysed from 55 eyes. Dice scores were compared between the three methods (RANSAC-SC/OR CNN/OR CNN + DC): 0.3341/0.4665/4784 for fold 1-2 and 0.3315/0.4494/4596 for fold 2-1 in composite images. The images composed using the OR CNN + DC have a median rating of 4 (out of 5) versus 2 using RANSAC-SC. The odds of getting a higher grading level are 4.8 times higher using our OR CNN + DC than RANSAC-SC ($p < 0.0001$).

CONCLUSION: Peripheral retinal vessel alignment performed better using our AI algorithm than RANSAC-SC. This may help improve co-localizing retinal anatomy and pathology with our algorithm.

Eye (2024) 38:1189–1195; <https://doi.org/10.1038/s41433-023-02868-3>

INTRODUCTION

In retinal practice, aside from fundus examination, diagnoses have been reliant on the use of digital imaging as most retinal imaging modalities are non-invasive and have fast acquisition, not only in 2D but also 3D [1]. The clinical fundus examination is currently used in conjunction with other wide-field imaging in patients with retinal diseases such as diabetes, sickle cell, tumours, peripheral choroidal neovascularization, and many more. The common modalities used are Fluorescein Angiography (FA), Indocyanine Green Angiography (ICGA), Fundus Autofluorescence (FAF), and Optical Coherence Tomography (OCT). Multimodal imaging has been increasingly popular in ophthalmology, particularly in the retina. The images that these devices provide are critical in the decision-making of clinicians for the diagnosis and management of certain ocular diseases. A study by the DRCR Retina Network group noted an association between predominantly peripheral lesions and the risk of worsening diabetic retinopathy using ultra-wide field imaging [2]. The Heidelberg Spectralis (Heidelberg Engineering, Heidelberg, Germany) uses a confocal scanning laser ophthalmoscope (cSLO) and captures images at three wavelengths:

short wavelength (486 nm), which captures the surface of the retina, medium wavelength (518 nm) which captures the retinal vascular and inner retinal layers, and long wavelength (815 nm) which captures the RPE and other deep layers [3]. The Optos devices (Optos plc, Dunfermline, United Kingdom), on the other hand, use a scanning laser ophthalmoscope (SLO) and capture images at two wavelengths: red laser (635 nm) and green laser (532 nm) [4]. These two imaging modalities are commonly used in the retinal clinic to provide insight into the patient's retina status and document certain pathologies. A study by Muftuoglu et al. [5] concluded that the multicolour (MC) imaging from Heidelberg Spectralis provided superior epiretinal membrane detection compared to conventional fundus photo imaging, which was primarily attributed to the green wavelength of the MC imaging. Commonly, retinal imaging can be performed using a scanning laser ophthalmoscope (SLO) at a wider angle, but most of these machines do not have a built-in OCT. The standard of care in OCT is typically confocal SLO (cSLO) imaging. However, different instruments have different wavelengths, hence different information. Certain instruments are better steered so that if there are peripheral

¹Jacobs Retina Center, University of California, San Diego, CA, USA. ²Viterbi Family Department of Ophthalmology and Shiley Eye Institute, University of California, San Diego, CA, USA. ³Division of Ophthalmology Informatics and Data Science, Viterbi Family Department of Ophthalmology and Shiley Eye Institute, University of California, San Diego, CA, USA. ⁴Department of Electrical and Computer Engineering, University of California, San Diego, CA, USA. ✉email: wrfreeman@health.ucsd.edu

Received: 1 June 2023 Revised: 7 November 2023 Accepted: 22 November 2023

Published online: 19 December 2023

lesions to be seen with an OCT, it can be done with a few manoeuvres. The standard clinical practice is to compare a lesion seen in two imaging modalities by bringing them up on a computer monitor and doing a side-by-side manual comparison. For example, suppose a choroidal tumour was imaged using a 55-degree montage of one MC imaging modality and another one from a colour photograph. In that case, these two composite images will be compared using a side-by-side method. This is time-consuming and may have inaccuracies in co-localizing regions of interest. An automated method to do this would be clinically beneficial and more rapid and accurate than manual methods.

Artificial intelligence (AI) is a general term implying the use of a computer to model intelligent behaviour with minimal human intervention. It is considered a branch of engineering in this generation. Its application in medicine has two branches: physical and virtual. The physical component is represented by robots to aid the physicians or surgeons and patients. The virtual component is represented by Machine or Deep Learning. Deep learning is represented by mathematical algorithms that improve learning through experience [6]. Deep learning has been introduced in other medical specialties like cardiology, dermatology, and neurology [7]. In ophthalmology, deep learning has been applied for medical imaging analysis in detecting various ocular conditions using mainly fundus photographs and optical coherence tomograph [8]. However, there are some limitations of deep learning-based models. If the models fail, unnecessary additional interventions or misdiagnosis may lead to poor outcomes [9]. That is why training the AI is essential to improve the analysis.

Cavichini et al. [10]. compared the conventional mathematical warping (Modality Independent Neighborhood Descriptor [MIND]) and a new AI method in aligning two types of retinal images taken in two different imaging modalities, i.e., colour fundus photographs and infrared scanning laser ophthalmoscope with narrow angles of view primarily of the posterior pole of the eye. They discovered their AI method superior to conventional warping at the posterior pole. This led us to use wider imaging modalities of different specifications to better understand the retinal periphery, as there has been no published literature regarding overlay of UWF images.

The purpose of our study was to compare a new AI method to conventional warping (Random Sample Consensus, Sample and Consensus set [RANSAC-SC]) in the alignment of retinal vessels at the retinal periphery using machines with different optical pathways, resolutions, and wavelengths: ultra-wide field (Optos) and wide-field (Heidelberg Engineering) images. We selected two imaging systems with different optics and wavelengths by performing an AI montage in the Heidelberg images to get the full view of the retina and overlay it with the ultra-wide field (Optos) image. Our group observed the quality of overlay in RANSAC-SC vs. our AI algorithms (Outlier Rejection Convolutional Neural Network [OR CNN] and OR CNN + Distortion Correction [OR CNN + DC]).

METHODS

Study participants

This study was approved by the Institutional Review Board of the University of California San Diego in California, USA. Data collection and analysis were conducted according to the Principles of the Declaration of Helsinki and complied with the Health Insurance Portability and Accountability Act of 1996. The patient's consent was obtained as per institution protocol, and all the data that were collected were anonymized.

Participants included in this study were consecutive patients that were seen at the Jacobs Retina Centre, University of California San Diego, from July 2021 until September 2021. Participants had retinal imaging done using Heidelberg HRA + OCT Spectralis System version 1.11.2.0 (Heidelberg Engineering, Heidelberg, Germany) and Optos Optomap Monaco P200TE software version 3.1.0.12 (Optos plc, Dunfermline, United Kingdom).

Image acquisition and processing

We collected the Ultra-wide field images (UWF), the wide field (WF) images, and generated the WF composite images in the following way: The UWF images were captured by the Optos ultra-wide field retinal imaging device with the colour Optomap image modality and the standard image view (200° internal angle = ~135° conventional nomenclature, single capture). For each eye, we took one UWF image. The WF images were captured using the Heidelberg Spectralis device (55° field-of-view [FOV]) with the following settings: Multicolour mode, ART = 15, High-speed mode and 3 × 3 matrix activated to capture nine fields of the retina. A total of nine WF images were captured in one eye.

A total of nine (9) 55-degree FOV images using the Heidelberg system and one single Optos image were captured in one eye. The Heidelberg steered images, when montaged, comprised a field of view ranging from 110–120 degrees. In this paper, we used the term ultra-wide field (UWF) for Optos images and wide-field (WF) for Heidelberg images.

To generate the WF composite image, we set the UWF image as a fixed reference image and then align each WF image independently to the UWF image. Finally, we fused all WF images from the same eye into one composite image. We applied Gaussian blur to the masks of the WF images to obtain smooth transitions at the edges of the WF overlaying areas.

Alignment algorithm

Our proposed pipeline for UWF-WF/UWF-Composite WF retina image alignment is an AI-based method and follows a feature-based registration procedure, including feature detection, description, matching as well as outlier rejection. The outlier rejection step is crucial in later transforming the retina image for accurate alignment. Therefore, we compared our AI implementation with the classic RANSAC-SC [11] method in this work. Compared with our previous work in Zhang et al. 2021 [12], in order to align WF images taken from different angles with a UWF image on different regions, we propose a distortion correction algorithm with 5 degrees of freedom for the WF camera to better eliminate the distortion at the outer region of the UWF images. Furthermore, in order to generate and accurately align the composite WF image, we propose a new warping strategy that first warp UWF towards WF to determine the best relative pose between two cameras and then compute the inverse transformation to warp all of the WF images towards the UWF modality.

In our proposed method (Fig. 1), we first extract vessel maps from both input UWF and WF images by two vessel segmentation convolutional neural network (CNN), respectively. The vessel segmentation CNN for the UWF modality is based on a model pre-trained on a public UWF dataset by Ding et al. [13]. Then, we set up a joint segmentation and deformable registration training procedure similar to the one published by Zhang et al. [14]. to train the vessel segmentation CNN for the WF modality. The segmentation knowledge stored in the pre-trained CNN for UWF can be migrated into the WF segmentation CNN by a deformable alignment network.

After vessel maps are extracted from both input images, we adopted a SuperPoint network [15] as the feature detection and description CNN to extract features from them. The SuperPoint Network has been pre-trained on a synthetic grayscale image dataset and will output key-point coordinates and their respective descriptions for each vessel map. Then, the detected features of both images are matched with each other based on bi-directional consensus, i.e., a feature in UWF should be the best match of a feature in the WF, and vice versa for the same pair of features.

Afterward, we used an outlier rejection CNN [16] to estimate inliers and their weights from the matched features. It takes as input the coordinates of all matched features and outputs weights for each. It has been pre-trained on a retinal registration dataset for Colour Fundus and Infrared images [17], and then fine-tuned on a small UWF-to-WF alignment dataset [12]. We specifically compared AI-based method with the previous RANSAC-SC method, which is specially designed to align two UWF images.

Finally, a transformation matrix was estimated from the matched features and weights by weighted least square (WLS) estimation. We applied second-order polynomial transformation to warp the WF images to align with the UWF image according to the matched inlier key points:

$$[u_i, v_i]^T = M_{poly}[m_i, n_i, m_i^2, n_i^2, m_i n_i, 1]^T$$

where $[u_i, v_i]$ and $[m_i, n_i]$ are keypoint locations in UWF and WF images, respectively. Here, M_{poly} is the 2×6 transformation matrix.

Due to the perspective distortion in the peripheral area of the UWF image, the above 2D registration method was unable to model the

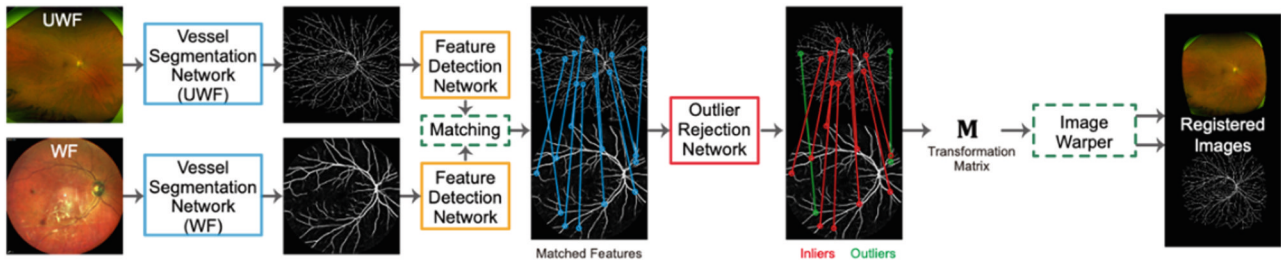


Fig. 1 Proposed one-way forward registration procedure. The proposed one-way forward registration procedure including CNNs for vessel segmentation, feature detection, description, and outlier rejection, respectively. CNN Convolutional Neural Network.

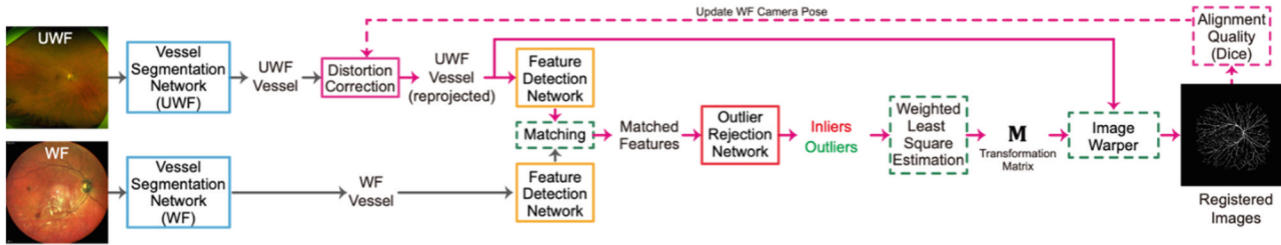


Fig. 2 The proposed iterative 3D camera pose optimization process. We first reproject the UWF using the distortion correction function based on the current estimate of WF camera pose, then estimate the transformation model based on the WF and the reprojected UWF images to warp the reprojected UWF image again. Next, compute the alignment quality (Dice) and finally update the WF camera pose for the next iteration. UWF ultra-wide field, WF wide-field.

distortion nor perfectly align the two images, even given perfect feature matchings. Therefore, after the CNNs have been set up in the forward 2D registration procedure, we proposed a distortion correction algorithm based on iterative 3D camera pose optimization to incorporate 3D eyeball shape information into the alignment process to improve alignment quality, as shown in Fig. 2. After the distortion correction algorithm, the two images can share similar levels of distortions which can be more easily modelled by 2D-to-2D transformation matrices. As shown in Fig. S1, we followed the assumptions of stereographic projection and spherical eyeball for UWF images, projected the 2D WF image grid onto a 3D sphere, and re-projected it onto the 2D imaging plane of the UWF image based on the WF camera pose. The camera pose is determined by 5 extrinsic parameters of the camera:

$$x_{cam} : (x, y, z); R_{cam} : (\theta_x, \theta_y)$$

where x, y, z are positions in 3D in the coordinate shown in supplemental Fig. S1, θ_x, θ_y are rotations of the camera in x and y directions. To be more specific, the grid (m, n) of the WF image is first converted to the 3D world coordinate (m', n', f') and projected onto the 3D retina surface (m'', n'', z) according to the camera pose. Then the 3D coordinates on eyeball are projected to the UWF image plane and converted to UWF pixel coordinate (m_p, n_p) . The UWF image can be warped to the same distortion level with the WF image according to the difference between the original and the projected grid:

$$I_{uwf-warped} = STN(I_{uwf}, (m_p, n_p) - (m, p))$$

The 3D eyeball is assumed to be a sphere with radius one, and the UWF camera is placed at the centre of cornea, i.e., $(0, 0, -1)$ in the world coordinate. The UWF image plane is $z = 1$ and the WF image plane's centre is also assumed to be on $z = 1$ while its orientation is determined by the WF camera pose. The distortion correction algorithm and the CNNs (2D pipeline) were concatenated into an iterative optimization process (Fig. S2), where we search for an optimized WF camera pose based on the alignment performance between the WF image and the warped UWF image.

Then, we used the optimized camera pose and the same 3D projection model to inversely warp the WF image and project the key point locations on it to the UWF image plane. Finally, the 2D alignment matrix was calculated from the UWF key points and the projected WF key points and

warped the WF image to UWF again in 2D. Single warped WF images were then assembled to the composite image to be aligned with UWF image.

Composite image generation algorithm

For each eye, there is one UWF image with 8–9 WF images. After the WF images and the corresponding binary mask were warped to the UWF image through the 3D projection and 2D alignment, we summed all the WF images from one eye along with the masks. The masks' boundary was smoothed by Gaussian blurring to get a smoother transition between WF images in the final composite images. Finally, the summed image was divided by the summed mask to generate the composite image. Figure S3 shows the registration of nine 55-degree FOV WF images (A and B) with corresponding UWF image (C).

Implementation details

The UWF images have pixel resolution of 4000×4000 . The resolution of the original WF images is 768×768 before montaging and varies after montaging. The images were later expanded to the same size as the UWF images (interpolated) for wrapping and composite image generation. Our dataset (446 pairs) was randomly separated into set 1 (237 pairs) and set 2 (209 pairs). For fold 1-2, we tested our algorithm on set 1 and train on set 2. For fold 2-1, we tested set 2 and train on set 1.

The segmentation network was pretrained on colour fundus images [17] and fine-tuned on our dataset. We used learning rate $1e-3$ and trained for 2000 epochs. The key point detection network (Superpoint) is pretrained on a hybrid dataset with colour fundus and infrared reflectance retina images in the same way as previously done [17]. For the outlier rejection network, we used learning rate $1e-4$ and trained for 1000 epochs. We used Adam optimizer for both the segmentation and outlier rejection network training.

Grading method

We used a modified version of Dice score in this paper, which is defined as

$$Dice(I_1, I_2) = \frac{2 \cdot \sum pixelwise_{min}(I_1, I_2)}{\sum I_1 + \sum I_2}$$

where I_1 and I_2 are two vessel maps in range $[0,1]$.

Each overlaid image was presented as a checkerboard composed of alternating Heidelberg multicolour and Optos pseudocolor images. To avoid bias in terms of decision-making from the Likert 5 scale using the

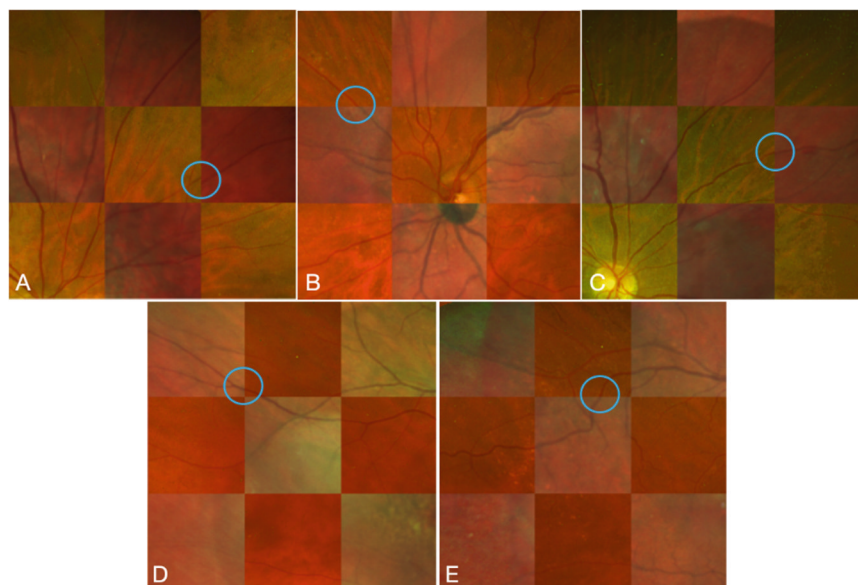


Fig. 3 Sample grading evaluation of eyes analysed. Grade 1 (A) = > 1 vessel width difference, Grade 2 (B) = $> \frac{1}{3}$ vessel width difference, Grade 3 (C) = $\frac{1}{3} - \frac{1}{2}$ vessel width difference, Grade 4 (D) = $\frac{1}{3}$ or less vessel width difference, Grade 5 (E) = perfect alignment.

generated algorithms RANSAC, OR CNN, or OR CNN + DC, the alignment of the peripheral vessels was graded independently by a masked grader (AL) from the overlaid images. For each overlaid image, an unmasked reviewer (FK) assigned a specific box from the checkerboard with a visible vessel in each adjacent box closest to the optic nerve, which was located to the periphery outside the vascular arcades. The unmasked reviewer located these points at the superotemporal, superonasal, inferotemporal, and inferonasal periphery. When there were multiple vessels seen per box, the unmasked grader assigned a specific vessel to grade. The alignment of vessels was based on the grading from the previous literature by Cavichini, et al. [10]. Grade 1 = > 1 vessel width difference, Grade 2 = $> \frac{1}{2}$ vessel width difference, Grade 3 = $\frac{1}{3} - \frac{1}{2}$ vessel width difference, Grade 4 = $\frac{1}{3}$ or less vessel width difference, Grade 5 = perfect alignment (Fig. 3). The masked graders and unmasked reviewer viewed the images in one high-quality widescreen monitor for uniformity. The results from the masked graders were tabulated and analysed by the unmasked graders (FK, LC).

Statistical analysis

Each composite image interest was graded by the same retina specialist three times using three different composing algorithms. The goodness of retinal vessel alignment at the peripheral retinal vessel arcade was graded into Likert 5 scales. Scales of vessel alignment goodness were compared among three composing algorithms by modelling the probabilities of having a higher scale value within an ordinary generalized estimating equation (GEE). All statistical analysis was performed using SAS software version 9.4.

RESULTS

We computed the Dice scores between the UWF and the warped WF images based on their respective vessel maps. When using the independent WF images (no composite), the Dice scores are 0.3407/0.4950/0.5133 for fold 1-2 and 0.3343/0.4654/0.4831 for fold 2-1 (RANSAC-SC/OR CNN/OR CNN + DC). When considering the WF composite images, the Dice scores are 0.3341/0.4665/0.4784 for fold 1-2 and 0.3315/0.4494/0.4596 for fold 2-1 (RANSAC-SC/OR CNN/OR CNN + DC). (Table 1) Our distortion correction algorithm requires all the previous steps to be deterministic so that we can search for the best camera extrinsic parameters under a deterministic 3D relationship. However, the result generated by RANSAC is random, bringing major noise to the algorithm. Therefore, we claim that our distortion correction algorithm is not suitable for the RANSAC method.

Table 1. Experiment results from different methods.

Dice	Single Image		Composite Image	
	Set1	Set2	Set1	Set2
RANSAC-SC [11]	0.3407	0.3343	0.3341	0.3315
OR CNN [14]	0.4950	0.4654	0.4665	0.4494
OR CNN + DC	0.5133	0.4831	0.4784	0.4596

RANSAC-SC Random Sample Consensus, Sample and Consensus sets, OR CNN Outlier Rejection Convolutional Neural Network, OR CNN + DC Outlier Rejection Convolutional Neural Network + Distortion Correction.

Two hundred and twenty boxes from 55 eyes were analysed per algorithm. This yielded a total of 660 boxes. GEE modelling demonstrated that the image composed using our AI algorithm OR CNN or OR CNN + DC has significantly better retinal vessel alignment than the images composed by the RANSAC-SC (Odds ratio 3.6, $p < 0.0001$; Odds ratio 4.8, $p < 0.0001$, respectively). For comparison of AI algorithm OR CNN with OR CNN + DC, OR CNN + DC is superior (Odds ratio 1.3, $p = 0.0013$). A sample comparison between our AI algorithms and RANSAC-SC is shown in Fig. 4. Regarding retinal vessel alignment using the Likert 5 grading scales, the images composed using our algorithm OR CNN + DC have a median rating of 4 (out of 5). In contrast, images composed using RANSAC-SC have a median rating of 2. This means that the alignment of retinal vessels was highest using our algorithm (5 being perfect alignment).

DISCUSSION

Artificial intelligence in the retina has been increasingly popular worldwide. In 2018, Stevenson et al. [18]. used AI to implement a photo-screening pathway for different retinal diseases. Other groups [19, 20] used deep learning systems, which accurately distinguished retinal pathologies in real time. Aligning the vascular landmarks between two imaging modalities will be important to accurately co-localize different retinal pathologies, particularly in the periphery of the retina.

Our group used these two imaging modalities since wide-angle imaging is also becoming more popular as peripheral

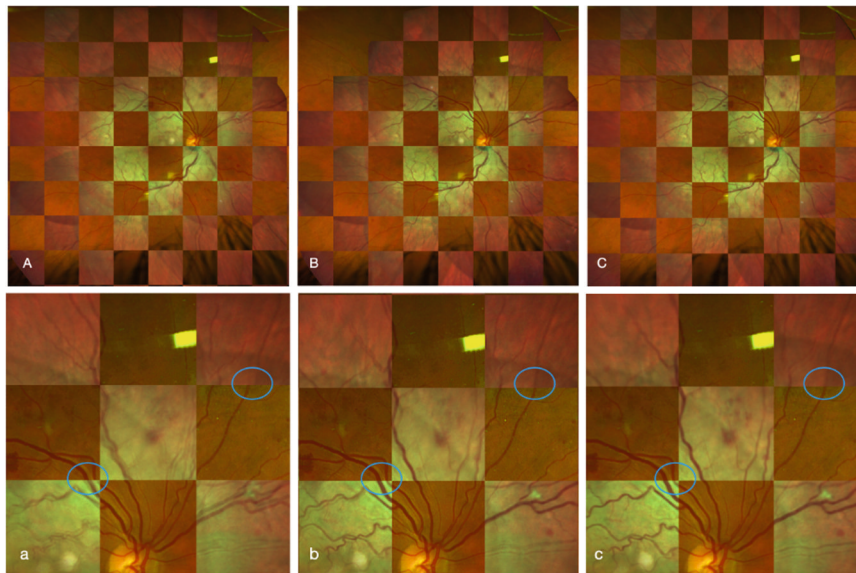


Fig. 4 Sample Heidelberg and Optos checkerboard overlay image. Sample Heidelberg and Optos checkerboard overlay image of RANSAC-SC (A), OR CNN (B), and OR CNN + DC (C) with its corresponding magnified image of alignment (a, b, and c, respectively). Images using the AI algorithms OR CNN and OR CNN + DC showed more aligned images than RANSAC-SC, as seen from the encircled vessels. RANSAC-SC Random Sample Consensus, Sample and Consensus sets, OR CNN Outlier Rejection Convolutional Neural Network, OR CNN + DC Outlier Rejection Convolutional Neural Network + Distortion Correction.

retinal or choroidal lesions may also be noted that may warrant further evaluation. Visualization of retinal or choroidal lesions may be different using different imaging modalities. A study from our group by Muftuoglu et al. [21]. compared the characteristics of retinal and choroidal lesions using colour fundus photography and multicolour imaging. We found that the infrared light of the MC imaging underestimates the extent of choroidal lesions by 33%. An innovation of the present study shows that using AI, we can accurately align and overlay images from different types of wide-angle imaging cameras to evaluate lesions outside the posterior pole of the retina. This will be useful when different instruments with different specifications are used.

The ability to colocalize lesions will be an important part of the management. For both natural and medical images, there are always mismatched features that will affect the registration performance. Both the AI model and RANSAC are trying to determine those mismatched features (i.e., outliers) and estimate the transformation matrix using the reliable features (i.e., inliers), yet the AI model has better capability in finding outliers because it can be trained and adapted to the data distribution of retinal images. While RANSAC is a fully randomized method independent of data, its performance is worse. Although the Optos camera captures an approximately 135-degree ultra-wide field image (135° external is the conventional nomenclature and is equal to 200° internal angle measured from the centre of the eye) [22], the ellipsoid mirror provides a different aspect ratio in the periphery than in the centre. Hence, this promotes minimal distortion along the periphery [23]. Our goal was to align the retinal peripheral images captured from the Heidelberg Spectralis to the Optos images using the retinal vasculature as a landmark using our algorithm and compare it with the conventional mathematical warping.

RANSAC is one of the most well-known algorithms because of its robustness and simplicity in solving geometric estimation problems in datasets containing outliers [24, 25]. Our algorithm provided a better alignment in the retinal vasculature than the conventional warping. When comparing the different algorithms, dice coefficients showed better scores using our algorithm; there

was mild improvement when OR CNN was used and a vast improvement when DC was added. When retinal vessel alignment scoring was performed, there was a much greater improvement in retinal vessel alignment noted. The clinical scoring we used focused on the peripheral retina, which showed better improvement when OR CNN + DC was used. The checkerboard presented in our study provided certain number of points per human checker, i.e., certain points per retinal image to check, and our results are highly statistically significant.

Theoretically, the centre area (macula) will always have better alignment than the peripheral area because the centre area does not have (perceivable) distortion between the two modalities to correct. Our study evaluated the peripheral retinal vessel alignment using our AI algorithm (OR CNN + DC), which showed superiority over the conventional method (RANSAC-SC). This could help identify lesion sizes (e.g., melanoma, nevus) and monitor progression to understand the pathology better using two imaging modalities. Our method should aid in a better comparison of lesions in the anterior retina. This can be the first step toward further analysis of retinal peripheral imaging, as other imaging modalities may provide insight into retinal architecture. A better understanding of other pathologies by overlaying an optical coherence tomography (OCT) to an algorithm like ours may properly co-localize lesions or abnormalities to the extent that close monitoring or treatment may be warranted. The distortion correction method proposed previously [12] was designed for the case where MC images are in the centre of the fovea. In that case, we assume the relative camera pose (UWF to MC) is mainly different in the distance to the eye while they roughly remain on the same axis. Therefore, we only use one parameter to model the difference between the image-capturing model for both modalities. Regarding the distortion correction model of the present study, since more MC images are in the peripheral retina, we assumed that there are more variations in the relative camera pose. Therefore, we used five parameters to describe the camera pose difference, which can cover more perspective distortions. Ideally, there are six parameters to describe the relative camera pose for perspective projection. We left one dimension, i.e.,

rotation, to be solved by the following 2D image warping process. Traditional feature descriptors and the work of DeTone et al. [15], were designed on natural images of similar appearance. However, the variations in modalities will reduce the reliability of their results. Therefore, we transformed multi-modal retinal images into the same modality, i.e., vessel maps, for the following feature detection, increasing the reliability of feature detection results.

To the best of our knowledge, this is the first study to perform an overlay of multicolour wide-field imaging to a pseudocolour ultra-wide field imaging as these two imaging modalities are more commonly used as they are convenient and fast, which lessens the patient burden, especially for uncooperative or unstable patients.

In conclusion, our AI algorithm improved vessel alignment in the retinal periphery compared to RANSAC-SC. This may aid in the future analysis of peripheral changes or pathologies and will allow precise alignment and overlay of peripheral retinal images using a variety of different devices. Other current imaging modalities – adaptive optics, optical coherence tomography/angiography, fluorescein angiography, microperimetry – may benefit from localizing retinal anatomy and pathologies.

SUMMARY

What was known before

- Wide-angle imaging, as well as high-resolution posterior pole imaging, are increasingly important in understanding and treating retinal disease
- It is difficult to precisely register and compare such images taken on different retinal imaging platforms

What this study adds

- This study shows that with the use of artificial intelligence, such images can be precisely overlaid and compared to better understand pathology across the entire retina

DATA AVAILABILITY

All data in the current study are available from the corresponding author (WRF) upon reasonable request.

REFERENCES

- Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina. *Prog Retin Eye Res.* 2018;67:1–29.
- Marcus DM, Silva PS, Liu D, Aiello LP, Antoszyk A, Elman M, et al. Association of predominantly peripheral lesions on ultra-widefield imaging and the risk of diabetic retinopathy worsening over time. *JAMA Ophthalmol.* 2022;140:946–54
- Heidelberg EngineeringSpectralis H. *MultiColor ATLAS.*
- User Guide P200TE Optos OCT *Ophthalmoscope.*; 2019. www.optos.com/patents.
- Kilic Muftuoglu I, Bartsch DU, Bartselli G, Gaber R, Nezgoda J, Freeman WR. Visualization of macular pucker by multicolor scanning laser imaging. *Retina* 2018;38:352–8.
- Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism* 2017;69:S36–S40.
- Date RC, Jesudasan SJ, Weng CY. Applications of deep learning and artificial intelligence in retina. 2019. <http://journals.lww.com/internat-ophthalmology>.
- Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol.* 2019;103:167–75.
- Korot E, Wagner SK, Faes L, Liu X, Huemer J, Ferraz D, et al. Will AI Replace Ophthalmologists? *Transl Vis Sci Technol.* 2020;9:2. <https://doi.org/10.1167/tvst.9.2.2>. Erratum in: *Transl Vis Sci Technol.* 2021;10:6

- Cavichini M, An C, Bartsch DUG, Jhingan M, Amador-Pattaroyo M, Long CP, et al. Artificial intelligence for automated overlay of fundus camera and scanning laser ophthalmoscope images. *Transl Vis Sci Technol.* 2020;9:1–10.
- Ding L, Kang TD, Kuriyan AE, Ramchandran RS, Wykoff CC, Sharma G. Combining feature correspondence with parametric chamfer alignment: hybrid two-stage registration for ultra-widefield retinal images. *IEEE Trans Biomed Eng.* 2023;70:523–32.
- Zhang J, Wang Y, Bartsch DUG, Freeman WR, Nguyen TQ, An C. Perspective distortion correction for multi-modal registration between ultra-widefield and narrow-angle retinal images. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS.* Institute of Electrical and Electronics Engineers Inc. 4086-91 (2021).
- Ding L, Kuriyan AE, Ramchandran RS, Wykoff CC, Sharma G. Weakly-supervised vessel detection in ultra-widefield fundus photography via iterative multi-modal registration and learning. *IEEE Trans Med Imaging.* 2021;40:2748–58.
- Zhang J, Wang Y, Dai J, Cavichini M, Bartsch DU, Freeman W, et al. Two-step registration on multi-modal retinal images via deep neural networks. *IEEE Trans Image Process.* 2022;31:823–38.
- DeTone D, Malisiewicz T, Rabinovich A. SuperPoint: Self-Supervised Interest Point Detection and Description. pp. 337–33712 (2018). <https://doi.org/10.1109/CVPRW.2018.00060>.
- Yi KM, Trulls E, Ono Y, Lepetit V, Salzmann M, Fua P. Learning to Find Good Correspondences, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 2666–74 (2018). <https://doi.org/10.1109/CVPR.2018.00282>.
- Wang Y, Zhang J, Cavichini M, Bartsch DU, Freeman W, Nguyen T, et al. Robust content-adaptive global registration for multimodal retinal images using weakly supervised deep-learning framework. *IEEE Trans Image Process.* 2021;30:3167–78.
- Stevenson CH, Hong SC, Ogbuehi KC. Development of an artificial intelligence system to classify pathology and clinical features on retinal fundus images. *Clin Exp Ophthalmol.* 2019;47:484–9.
- Dong L, He W, Zhang R, Ge Z, Wang YX, Zhou J, et al. Artificial Intelligence for Screening of Multiple Retinal and Optic Nerve Diseases. *JAMA Netw Open.* 2022;5:e229960.
- Lin D, Xiong J, Liu C, Zhao L, Li Z, Yu S, et al. Application of comprehensive artificial intelligence retinal expert (CARE) system: a national real-world evidence study. *Lancet Digit Health.* 2021;3:e486–e495.
- Muftuoglu IK, Gaber R, Bartsch DUG, Meshi A, Goldbaum M, Freeman W. Comparison of conventional fundus photography and multicolor imaging in choroidal or retinal lesions. *Graefes Arch Clin Exp Ophthalmol.* 2018;256:643–9.
- Espina MP, Arcinue CA, Ma F, Camacho N, Bartsch DUG, Freeman WR. Analysis of a confocal scanning laser ophthalmoscope noncontact ultra-wide field lens system in retinal and choroidal disease. *Retina* 2017;35:2664–8.
- Kato Y, Inoue M, Hirakata A. Quantitative comparisons of ultra-widefield images of model eye obtained with Optos® 200Tx and Optos® California. *BMC Ophthalmol.* 19 (2019).
- Fotouhi M, Hekmatian H, Kashani-Nezhad MA, Kasaei S. SC-RANSAC: spatial consistency on RANSAC. *Multimed Tools Appl.* 2019;78:9429–61.
- 2009 IEEE 12th International Conference on Computer Vision (ICCV): Kyoto, Japan, September 27–October 4, 2009. *IEEE* (2009).

AUTHOR CONTRIBUTIONS

FGPK—designed the study, acquired and analysed the data, interpreted the data, drafted and revised the manuscript; MC—acquired the data; JZ—acquired and analysed the data, drafted the manuscript; BW—acquired and analysed the data, revised the manuscript; AL—analysed the data; AH—analysed the data; TN—designed the study, analysed the data, revised the manuscript; CA—designed the study, analysed the data; DUB—acquired and analysed the data, revised the manuscript; LC—analysed the data and revised the manuscript; WRF—conceptualized and designed the study, analysed the data, revised the manuscript. All authors listed have contributed to the work and approved the final version of the manuscript.

FUNDING

NIH grant OT2OD032644 (Kalaw, Heinke), Multimodal Retina Image Alignment and Applications R01EY033847 (Nguyen and Freeman), UCSD Vision Research Centre Core Grant from the National Eye Institute P30EY022589, NIH grant R01EY016323 (Bartsch), an unrestricted grant from Research to Prevent Blindness, NY (Freeman), and unrestricted funds from the UCSD Jacobs Retina Centre.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41433-023-02868-3>.

Correspondence and requests for materials should be addressed to William R. Freeman.

Reprints and permission 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) 2023