

Artificial Intelligence for Automated Overlay of Fundus Camera and Scanning Laser Ophthalmoscope Images

Melina Cavichini^{1,2}, Cheolhong An³, Dirk-Uwe G. Bartsch¹, Mahima Jhingan^{1,4}, Manuel J. Amador-Patarroyo^{1,5}, Christopher P. Long⁶, Junkang Zhang³, Yiqian Wang³, Alison X. Chan⁶, Samantha Madala⁶, Truong Nguyen³, and William R. Freeman¹

¹ Jacobs Retina Center, Shiley Eye Institute, University of California San Diego, La Jolla, CA, USA

² Departamento de Oftalmologia, Faculdade de Medicina do ABC, Santo Andre, Brazil

³ Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA, USA

⁴ Aravind Eye Hospital, Madurai, India

⁵ Escuela Superior de Oftalmologia, Instituto Barraquer de America, Bogota, Colombia

⁶ University of California San Diego School of Medicine, La Jolla, CA, USA

Correspondence: William R. Freeman, Jacobs Retina Center, Shiley Eye Institute, University of California San Diego, 9415 Campus Point Drive 0946, La Jolla, 92093 CA, USA. e-mail: wrfreeman@ucsd.edu

Received: April 12, 2020

Accepted: September 23, 2020

Published: October 20, 2020

Keywords: artificial intelligence; multimodal images; retina; imaging; diagnosis

Citation: Cavichini M, An C, Bartsch D-UG, Jhingan M, Amador-Patarroyo MJ, Long CP, Zhang J, Wang Y, Chan AX, Madala S, Nguyen T, Freeman WR. Artificial intelligence for automated overlay of fundus camera and scanning laser ophthalmoscope images. *Trans Vis Sci Tech*. 2020;9(2):56. <https://doi.org/10.1167/tvst.9.2.56>

Purpose: The purpose of this study was to evaluate the ability to align two types of retinal images taken on different platforms; color fundus (CF) photographs and infrared scanning laser ophthalmoscope (IR SLO) images using mathematical warping and artificial intelligence (AI).

Methods: We collected 109 matched pairs of CF and IR SLO images. An AI algorithm utilizing two separate networks was developed. A style transfer network (STN) was used to segment vessel structures. A registration network was used to align the segmented images to each. Neither network used a ground truth dataset. A conventional image warping algorithm was used as a control. Software displayed image pairs as a 5 × 5 checkerboard grid composed of alternating subimages. This technique permitted vessel alignment determination by human observers and 5 masked graders evaluated alignment by the AI and conventional warping in 25 fields for each image.

Results: Our new AI method was superior to conventional warping at generating vessel alignment as judged by masked human graders ($P < 0.0001$). The average number of good/excellent matches increased from 90.5% to 94.4% with AI method.

Conclusions: AI permitted a more accurate overlay of CF and IR SLO images than conventional mathematical warping. This is a first step toward developing an AI that could allow overlay of all types of fundus images by utilizing vascular landmarks.

Translational Relevance: The ability to align and overlay imaging data from multiple instruments and manufacturers will permit better analysis of this complex data helping understand disease and predict treatment.

Introduction

As retinal treatments advance and imaging becomes more important, it will be critical to be able to scientifically analyze and interpret a large amount of information from different instruments, manufacturers, and diagnostic sources.¹ Many investigators have also found that imaging with different instruments or optics is useful in improving diagnosis and prognostic information.^{2–4} These clinical tools, however, have

multiple models, generations of software, and device specific algorithms used to output data. Ideally, all of this information could be organized by aligning such data by retinal location, which could then be interpreted using artificial intelligence (AI).⁵ It will be important for an AI agent to overlay data from a given retinal region that is procured from different imaging and function analysis instruments.

Previous studies have used AI as a multimodal registration method.^{6–9} Hervella et al. proposed a hybrid methodology for the multimodal registration of color

fundus retinal imaging and fluorescein angiography data that exploits the presence of the retinal vascular tree in retinal images.⁶ Mahapatra et al. applied the generative adversarial network to register multimodal images with the supervision of registration files, which are obtained from other conventional methods.⁷ However, in both studies, the overlay approach was limited to retinal images taken with the identical camera and the same field of view, just with different wavelengths (fluorescein angiography and color fundus images taken with a standard camera).

Additionally, AI has been used in analyzing single modality image analysis to categorize or detect disease,^{10–12} but there is no current method to co-localize and analyze multiple imaging and functional data. For this reason, as a preliminary step to applying AI to analyze multi-instrument imaging and functional studies, we attempted to overlay images from a scanning laser platform onto a fundus camera platform. These imaging platforms utilize different optical pathways as well as different types of illumination (scanning laser versus flood illumination). We chose to use an infrared scanning laser ophthalmoscope (IR SLO) image as a prototypical SLO image to overlay onto color fundus (CF). Photographs were taken with a fundus camera because such imaging is done on all patients undergoing optical coherence tomography (OCT) scans, also the optics and aspects ratio of infrared images are expected to be similar to and thus apply to autofluorescence (AF) or multicolor (MC) images taken with SLO so these results may be applicable to many types of images. We note that the SLO image is taken using different optics and instruments than is the CF image, so this appeared to be a good first step to determine if an AI agent can accomplish such overlaying by examining vessel locations. The novelty of this work is that we have conducted a rigorous, masked study of the performance of a novel AI algorithm for the alignment of multimodal retinal images. Our algorithm was able to perform image alignment without the need for a large set of manually annotated ground truth image sets.

Methods

This study was conducted according to the principles of the Helsinki Declaration. Institutional review board (IRB) approval was acquired from the University of California San Diego for the review and analysis of patient's data. The study complied with the Health Insurance Portability and Accountability Act of 1996.

Consecutive 50 degrees of diagonal field-of-view (FOV) CF images (TRC-50DX color fundus images, Topcon, Oakland, NJ) and 30 degrees × 30 degrees FOV (equal to 42 degrees diagonal) infrared images Scanning Laser Ophthalmoscope images (HRA + OCT Spectralis, Heidelberg Engineering, Heidelberg, Germany) were obtained between January 2017 and November 2018. We evaluated 1742 de-identified images from healthy eyes as well as eyes with retinal diseases, such as diabetic retinopathy, wet and dry macular degeneration, and retinovascular occlusion, in patients from our tertiary retina center (Jacobs Retina Center, Shiley Eye Institute, University of California San Diego, San Diego, California). The inclusion criteria were eyes with good quality images in CF photographs and IR SLO, taken on the same day.

We selected 1388 consecutive series of cases with good quality images taken with both a conventional fundus camera and IR SLO on the same day. One hundred thirty-eight images were excluded because they were not taken on the same day and pathology could have changed over time altering vessel position or focus plane, 216 images, which could not be evaluated properly due to poor focus, reflection, darkness, overexposure, or other artifacts made it difficult to identify vessels in one or both image types.

From the total of selected images, 1170 were used to train the AI, 667 healthy eyes and 503 eyes with retinal diseases and 218 different images were saved to be used after training to evaluate the overlay systems, 124 healthy eyes and 94 eyes with retinal disease. Two different overlaying algorithms were applied. One algorithm was a conventional warping - modality independent neighborhood descriptor (MIND) algorithm and the other an AI algorithm. Human graders masked to the method determined accuracy of overlay of the vessels. Both methods involved overlay of region images with cropping and rotations as needed to have both images on the same position.

The conventional alignment method or MIND proposed deformable multimodal registration based on human-engineered feature descriptor, which is based on the self-similarity of multimodal images.¹³ MIND was originally applied to register computed tomography and magnetic resonance imaging and was extended later to register retina images.¹⁴ This method uses mathematical warping algorithms to align vascular landmarks and is not an AI-based method.¹⁵ Such locally deformable registration can give a better estimation of transformations compared to affine registration but multimodal retinal image registration is still a challenge due to the modality dependent resolution, contrast, and luminosity variation between different modalities.^{14,16,17}

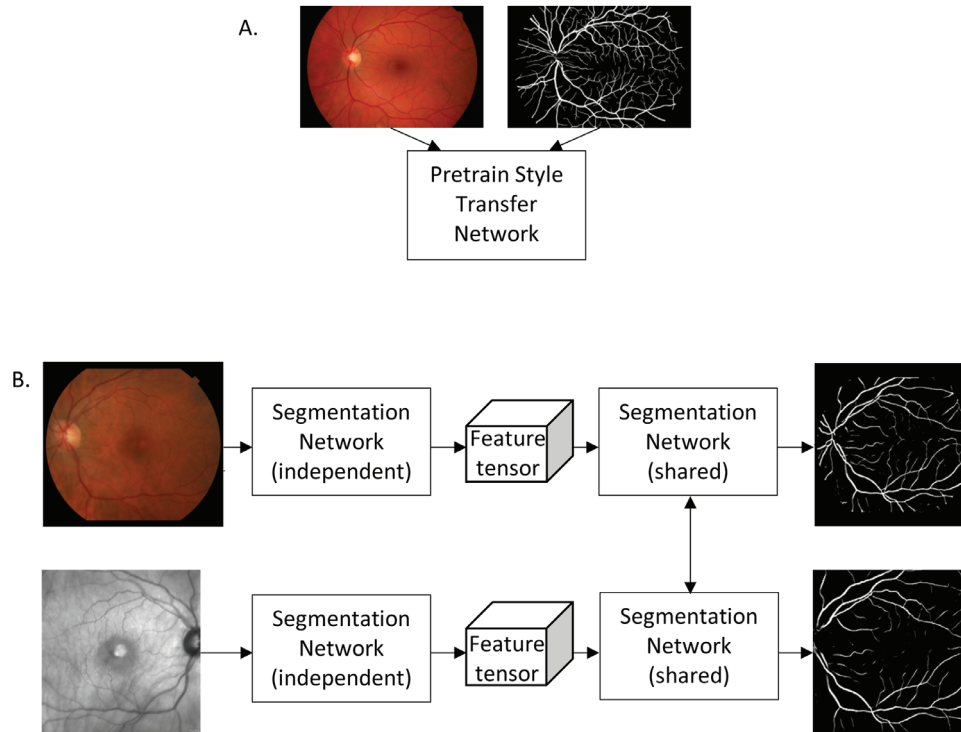


Figure 1. Diagram of the style transfer network (STN).¹⁸ (A) Shows the pretraining of the CNN using a single pair of fundus image and segmented vessel diagram. This image was not part of our dataset and was used to train our STN to detect tree-like structures, continuously stretching and branching vessel paths with decreasing width, etc. We only assume that both the style image and our retinal images share vessel structure styles. (B) Shows the application of the STN to our set of roughly aligned retinal images. Our network has an independent part and shared part. In the independent part the network calculates a feature tensor while removing the spatial information and only the summaries of styles is preserved. Multiple layers of increasing level of the network are used to detect patterns. In the shared part, the last layer of the network with sigmoid function is shared to guide the transform of the multimodal images into consistent representation of similar modalities. More details are found in our prior publication.¹⁸

Our AI overlay strategy consisted of a joint vessel segmentation and a deformable registration model based on the convolutional neural network because retina vessels are key landmarks even for different imaging modalities.^{18,19} The proposed learning scheme utilized two learning networks. First, a style transfer network was applied²⁰ to train a vessel segmentation without ground truth such that it would extract mutual patterns between multimodal retinal images to find good correspondences.²¹ We used one previously published vessel segmentation as initial training set and experimentally chose one vessel from the image to get the best performance. One segmented image, which is described as a style image, was used for all data (Fig. 1).¹⁸ The style transfer network transformed input retina images to target style images (segmented vessel images) as shown in Figure 1. The style transfer network (STN) uses a pre-trained convolutional neural network (CNN) to model the global vessel structure with an outside dataset (represented

by the image in Figure 1). The segmentation map was labeled by hand from the DRIVE dataset.²² This outside dataset was used as the style target. We only assumed that this style target and our retinal images share similar vessel structural styles (tree-like structure, continuously branching, and stretching vessels with decreasing width, for example). The STN has an independent and shared segmentation network. In the independent part, the network calculates a feature tensor while removing the spatial information and only the summaries of styles is preserved. Multiple layers of increasing levels of the network are used to detect patterns. In the shared part, the last layer of the network with sigmoid function is shared to guide the transform of the multimodal images into consistent representation of similar modalities. The deformable registration network was trained to find dense correspondence based on consistent vessel representations and wrapped image alignment.²³ The registration network provides the alignment information

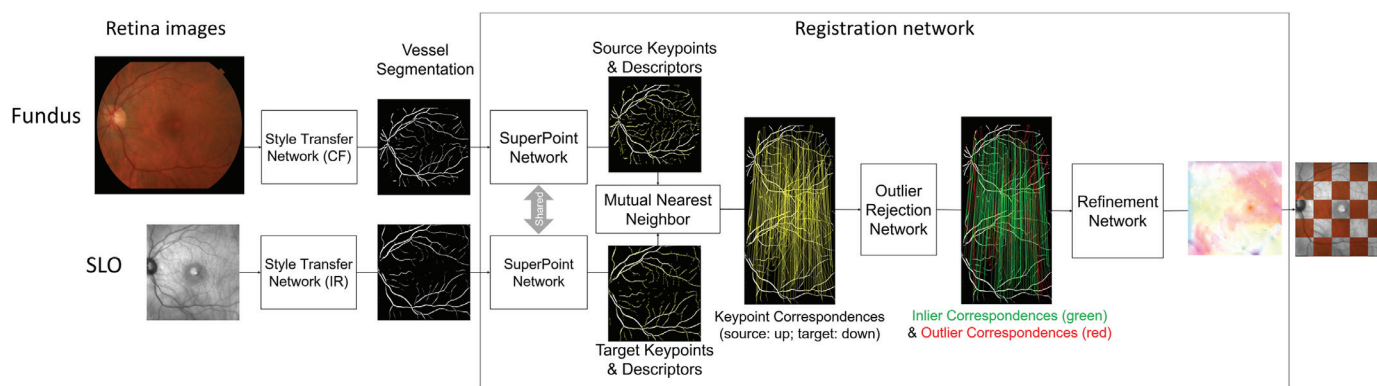


Figure 2. The structure of the proposed neural network to overlay multimodal retina images. The input images are the color fundus image and the SLO image shown on the left. The style transfer network (STN) is explained in more detail in Figure 1. The output of the STN is two vessel segmentations, as shown in the figure. The Registration Network consists of the super-point network³⁵ for key point detection and description, the outlier rejection network³⁶ for reliable matching points selection, and the refinement network¹⁸ for sub-pixel level adjustment. The super-point network determines the key points on the segmented vessel (denoted *yellow* points) and the corresponding descriptions. Next, key points of CF and ones of IR are matched with the nearest neighbor criterion, which is depicted as connection with *yellow lines*. More robust matching key points are derived with the outlier rejection network and inliers denoted as *green* connections are used for alignment. The refinement network provides sub-pixel level alignment information of which image represents the direction and magnitude of the localized image shift needed to achieve congruence between the source image (color fundus) and target image (SLO). The final image on the right shows the overlap of the color fundus image onto the SLO image in an arbitrary 5×5 square pattern.

from a source image (CF image) to a target image (SLO image) and aligned both segmented vessel images of two modalities and original retina images. The Registration Network consists of the super-point network²⁴ for key point detection and description, the outlier rejection network²⁵ for reliable matching points selection, and the refinement network¹⁸ for sub-pixel level adjustment. The registration network was also trained without any labeled data because it is impossible to obtain dense correspondences for retina images. These two networks were cascaded and trained via end-to-end learning in Figure 2. More details on the network or reports to replicate the results have previously been published for us.^{18,19} We used 109 datasets to train the algorithm, 20 datasets for validation and 89 datasets for testing.

A software was developed to show 436 evaluations, recorded as a result of 109 unique image pairs multiplied by 2 as block replicates, overlying both the techniques – with and without AI – each image was divided into 25 squares (5×5), which means a total of 5450 squares were compared. Each image was a checkerboard composed of alternating infrared and CF pictures. Each image was graded two times in each technique, and each individual square was graded based on the alignment of the vessels. Graders were masked because they could not identify which images were aligned by AI or MIND as they were presented in the same configuration.

The images were graded independently by two retina specialists and three medical students based on the longest vessel traversing the image zone (25 image zones or boxes per fundus photograph). First, each of the 5 graders scored 10 images and these images were subjected to the interclass correlation coefficient (ICC) using SPSS, and the ICC average among all the 5 graders was 0.903, which is considered excellent. The grading was performed in each zone by evaluating the vessel overlap in the area closest to the optic nerve. The alignment of the 2 images was graded 0 to 5, where 5 is a perfect alignment, 4 is less or equal to one-third the vessel width difference in continuity of the vessels, 3 is more than one third or equal to one-half the vessel width difference in continuity between the 2 vessels, 2 is more than one half and less than 1 vessel width difference in continuity between vessels, 1 is more than 1 vessel width difference in continuity between vessels, and 0 is ungradable due to absence of vessels (Fig. 3). For the analysis, we considered only regions where visible vessels were included. Grades 1 and 2 were considered a bad match, 3 reasonable, and 4 and 5 good/excellent matches.

After this result, a total of 5450 pairs of images zones were analyzed and compared (Fig. 4). We performed the Wilcoxon Signed Rank Test comparison between both methods using SPSS (IBM, version 26). Non-parametric statistics were used because of the categorical nature of the grading system.

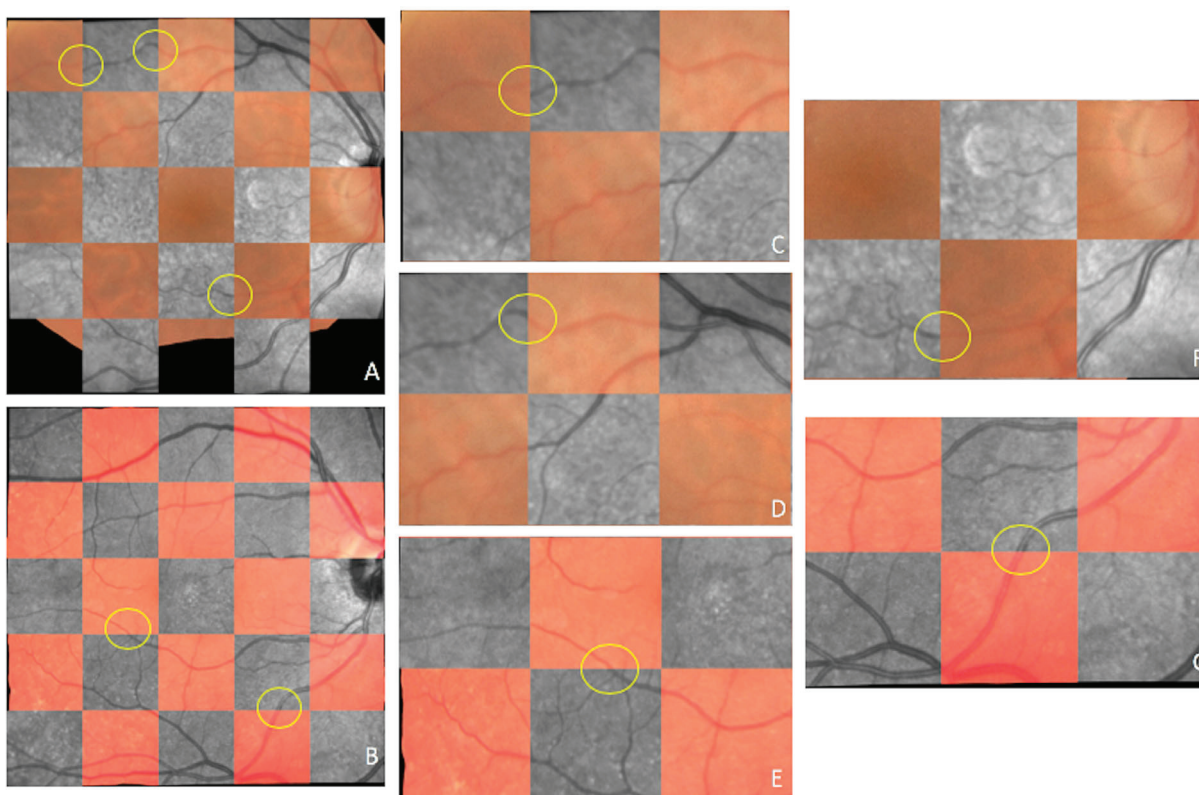


Figure 3. Examples of images analyzed using the checkerboard comparison method of grades 1 to 5 (0 ungradable, not shown). (A, B) The “mosaic,” overlying color fundus image and infrared SLO image. Each square was graded following the largest vessel closest to the optic nerve. (C, D, E, F, G) Are examples of classifications 1, 2, 3, 4, and 5, respectively, the yellow circles show the areas where the vessels’ alignment was scored in each square.

Results

We performed the Wilcoxon Signed Rank Test comparison between both methods. The AI overlay method was statistically significantly better than the conventional warping grading score as judged by masked grading by the experienced human graders ($Z = -8.467$, $P < 0.0001$).

There were 5,450 squares analyzed (25 squares per eye). The mean score of the conventional MIND procedure was 4.45 ± 1.228 and the mean score of the proposed new method was 4.58 ± 1.078 (Table). Even though the data are categorical ordinal and the proper statistical expression would be the median score, we decided to calculate the mean score and add this information to the table to show a difference. The statistical test found a highly statistical difference but the median scores of both procedures were identical at the value “5” and therefore meaningless. In particular, the number of bad matches was reduced by approximately 75% using the AI (proposed) agent

and there was also an increase in the proportion of good/excellent matches. In general, the assessors did not notice any systematic bias in images for regions that were consistently misaligned, although this was not formally assessed in this study. It was clear from the image alignment results that such alignment was nearly pixel to pixel using the AI (good/excellent matches) and the AI achieved this in 94.4% of cases, which was higher than the 90.5% achieved by the conventional warping method. Perhaps more importantly, the AI bad mismatch rate was 0.97% versus conventional warping of 3.8%. Thus, the AI can more closely overlay CF and IR SLO images analyzing lesions more precisely using the two modalities.

Discussion

The use of AI, in particular, deep learning, has been limited in retinal analytics but does show promise. A group in Germany analyzed predictors and visual

Table. Descriptive Statistics Wilcoxon Signed Rank Test ($Z = -8.467, P < 0.0001$)

Method	N	Mean	Standard Deviation	Bad Matches (1/2)	Reasonable Matches (3)	Good/Excellent Matches (4/5)
MIND (conventional warping)	5450	4.45	1.228	211/5450 = 3.84%	82/5450 = 1.5%	4935/5450 = 90.5%
AI	5450	4.58	1.078	53/5450 = 0.97%	36/5450 = 0.66%	5144/5450 = 94.4%
P value	$P < 0.0001$	Mean \pm SD 4.45 \pm 1.23 4.58 \pm 1.08	Grade 0 ^a 222 (4.1%) 217 (4.0%)	Grade 1 166 (3.0%) 46 (0.84%)	Grade 2 45 (0.84%) 7 (0.13%)	Grade 3 82 (1.5%) 36 (0.66%)
					Grade 4 899 (16.5%) 925 (17%)	Grade 5 4036 (74%) 4219 (77.4%)
MIND	5450					
AI	5450					
P value	$P < 0.0001$					

^aGrade 0 is ungradable due to absence of vessels.

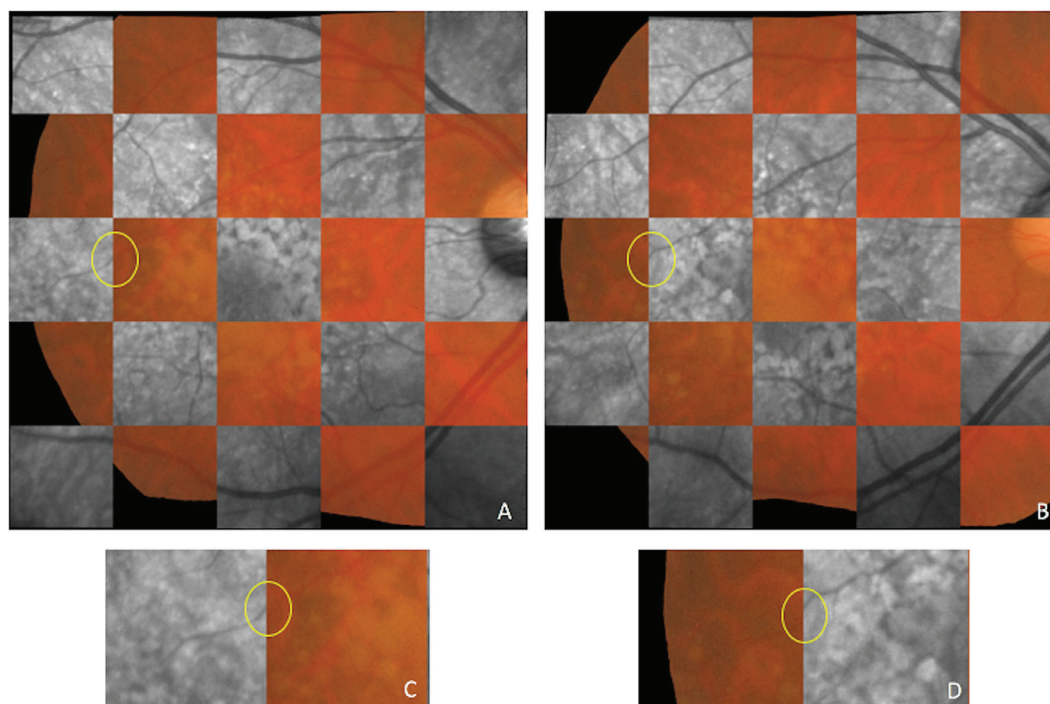


Figure 4. (A, B) The “mosaic,” overlaying CP and IR using the AI and the conventional (MIND) methods, respectively, in the same eye. On the border between squares 11 and 12, following the largest vessel closest to the optic nerve, it is possible to see the difference between both methods. (C) The border on the AI method, the vessel alignment classification (yellow circle) was 4, almost perfect. In (D) we can see the same border using the MIND method, the classification was 1, the poorest one.

outcomes of anti-vascular endothelial growth factor (VEGF) therapy and noted that analysis of raw imaging data would enhance predictive ability.⁹ Most analyses of retinal images evaluate only one type of imaging modality, a major problem in scientific rigor, and use human graders or AI algorithms analyzing only one modality, such as OCT layers.²⁶ As a preliminary step to using AI to overlay the plethora of different types of retinal images and functional tests, we evaluated overlay methods using AI and conventional warping algorithms from different imaging modalities, optics, and cameras. Our eventual goal is to be able to overlay multiple platforms. We are interested in this because numerous studies have suggested that analysis of OCT and other imaging and functional data may offer better ways to predict vision outcomes after choroidal neovascularization (CNV) treatment, but quantification of OCT raster scans, fluorescein leakage dynamics, volumes of retinal pigment epithelium (RPE) detachment, and subretinal fluid, fundus tessellation, and wide field angiography are difficult to perform without machine learning.^{27,28} In addition, our grading system was fine enough to detect vascular structures and misalignment down to 20 to 30 microns. This would encompass most retinal vascular abnormalities.

As a first step toward multi-instrument and modality image registration, we chose to develop an AI algorithm that would permit an overlapping of images from CF and scanning laser platforms. These are two completely different platforms for retinal imaging and do not readily overlay because of different optical pathways and light sources and illumination techniques. The FOV of the IR is 30×30 degrees and the CF is 50 degrees diagonal, our algorithm is very robust to different images conditions. Probably the main advantage of our approach is the ability to use CNN to alignment different retinal images without a ground truth (GT). There are very few GT databases available for training (DRIVE,²² and VARIA²⁹). Therefore, we wanted to develop an algorithm that does not require a GT database. In the absence of GT, we validated the results of the overlay with expert human observers using a grading system.

We chose to evaluate the ability to overlay CF images onto IR SLO because color imaging has been the standard for retinal evaluation for close to a century and does give the images that are the most similar to that of a clinical ophthalmoscopic examination. On the other hand, infrared images are a prototypical reference fundus image for SLO images and OCT scans and have similar optics and aspect ratio to AF

and multicolor scanning laser images. New SLO and other imaging systems allow wavelength and SLO selective imaging and are often done in combination with OCT. Confocal imaging is often used in SLO imaging to selectively image certain outer retinal or choroidal structures.³⁰ In addition, other imaging modalities, like AF imaging, utilize the principle of the property of AF to determine the size and activity of lesions.³¹ Different AF wavelengths may selectively image different photophores and or tissues as has been shown in Stargardt's disease, where infrared AF, which is commonly imaged with fundus cameras, picks up a larger size of the lesion as it reflects the AF from the RPE, as compared to the short wave AF imaging, which reflects the lesion at the level of the outer retina more than the RPE.³² Multicolor images recreate a CF image but differ in the ability to detect lesion features, particularly regarding retinal and choroidal pathologies.³³ Indeed, choroidal Nevi may appear larger and more prominent in fundus camera based near infrared imaging than SLO short wavelength AF imaging.^{34,35} We recognize that SLO imaging instruments are expensive but we chose to use them because it has the potential to deliver many types of useful imaging data. Furthermore, IR imaging of the retina is the standard for almost all OCT imaging instruments, including three of the four leading OCT manufacturer Heidelberg Engineering, Optovue, and Zeiss.

Our study shows the superior ability of AI as compared to conventional mathematical image warping programs to permit accurate overlaying and registration of images of the same fundus taken with two different imaging systems. The AI system was superior to conventional methods. We were careful to use a different data set to "train" the AI than what was used to evaluate it compared to both techniques. Because this has been demonstrated with two different systems (conventional wide-field flood camera fundus imaging and monochromatic scanning laser imaging) our results demonstrate the potential utility of AI in improving the problem of analysis of multimodal and multicamera (and functional) imaging in the field of retinal diseases, likely the modest but significant improvement in overlay by AI will become more important when widefield and/or more than two imaging modalities are analyzed. Future analytic techniques may allow the ability to simultaneously analyze angiography, OCT angiography, OCT, nerve fiber layer analysis data, microperimetry, wide-field imaging, and other techniques, such as adaptive optical imaging. Such analytics will improve our ability to better understand the parameters that best predict outcomes and help us understand retinal diseases. Features typically taken on SLO instruments, such as imaging of the

photoreceptor integrity, will potentially be able to be co-localized with multi wavelength AF images, OCT angiography, and conventional fluorescein or indocyanine green angiography and adaptive optical imaging. Once it is possible to co-localize structural imaging and functional imaging, such as SLO microperimetry,³⁶ such data can be combined with visual acuity, drug treatment, disease, and other information to help understand retinal disease better and also help predict outcomes to treatment. We recognize that our work was done in the central 30 degrees of the fundus where there is less distortion of images than when viewing the periphery. Further work will be needed to study peripheral retinal images and the ability to overlay those using different types of imaging techniques.

Acknowledgments

None of the authors had any financial/conflicting interests to disclose. The funding organizations had no role in the design or conduct of this research.

Supported in part by UCSD Vision Research Center Core Grant from the National Eye Institute P30EY022589, NIH grant R01EY016323 (DUB), an unrestricted grant from Research to Prevent Blindness, NY (W.R.F.) and unrestricted funds from the UCSD Jacobs Retina Center.

Disclosure: **M. Cavichini**, None; **C. An**, None; **D.-U.G. Bartsch**, None; **M. Jhingan**, None; **M.J. Amador-Patarroyo**, None; **C.P. Long**, None; **J. Zhang**, None; **Y. Wang**, None; **A.X. Chan**, None; **S. Madala**, None; **T. Nguyen**, None; **W.R. Freeman**, None

References

1. Kolar R, Kubecka L, Jan J. Registration and fusion of the autofluorescent and infrared retinal images. *Int J Biomed Imaging*. 2008;2008:513478.
2. Bandello F, Sacconi R, Querques L, Corbelli E, Cicinelli MV, Querques G. Recent advances in the management of dry age-related macular degeneration: a review. *F1000Res*. 2017;6:245.
3. Vidinova CN, Gouguchkova PT, Vidinov KN. Fundus autofluorescence in dry AMD - impact on disease progression. *Klin Monbl Augenheilkd*. 2013;230(11):1135–1141.
4. Fleckenstein M, Mitchell P, Freund KB, et al. The progression of geographic atrophy secondary to

- age-related macular degeneration. *Ophthalmology*. 2018;125(3):369–390.
5. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172(5):1122–1131.
 6. Hervella AS, Roucoa J, Novoa J, Ortega M. Multimodal registration of retinal images using domain-specific landmarks and vessel enhancement. *Procedia Computer Science*. 2018;126:97–104.
 7. Mahapatra D, Antony B, Sedai S, Garnavi R. Deformable medical image registration using generative adversarial networks. Presented at the: *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 1449–1453.
 8. Yoo TK, Choi JY, Seo JG, Ramasubramanian B, Selvaperumal S, Kim DW. The possibility of the combination of OCT and fundus images for improving the diagnostic accuracy of deep learning for age-related macular degeneration: a preliminary experiment. *Med Biol Eng Comput*. 2019;57(3):677–687.
 9. Miri MS, Abràmoff MD, Lee K, et al. Multimodal segmentation of optic disc and cup from SD-OCT and color fundus photographs using a machine-learning graph-based approach. *IEEE Trans Med Imaging*. 2015;34(9):1854–1866.
 10. Bernardes R, Lobo C, Cunha-Vaz JG. Multimodal macula mapping: a new approach to study diseases of the macula. *Surv Ophthalmol*. 2002;47:580–589.
 11. Arcadu F, Benmansour F, Maunz A, et al. Deep learning predicts OCT measures of diabetic macular thickening from color fundus photographs. *Invest Ophthalmol Vis Sci*. 2019;60(4):852–857.
 12. Fauw JD, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Med*. 2018;24(9):1342–1350.
 13. Heinrich MP, Jenkinson M, Bhushan M, et al. Mind: modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*. 2012;16(7):1423–1435.
 14. Li Z, Huang F, Zhang J, et al. Multi-modal and multivendor retina image registration. *Biomed Opt Express*. 2018;9(2):410–422.
 15. Heinrich MP, Jenkinson M, Bhushan M, et al. Mind: Modality independent neighborhood descriptor for multi-modal deformable registration. *Medical Image Analysis*. 2012;16(7):1423–1435.
 16. Liu C, Ma J, Ma Y, Huang J. Retinal image registration via feature-guided Gaussian mixture model. *Journal of the Optical Society of America*. 2016;33:1267.
 17. Miri MS, Abràmoff MD, Kwon YH, Garvin MK. Multimodal registration of SD-OCT volumes and fundus photographs using histograms of oriented gradients. *Biomedical Optics Express*. 2016;7:5252.
 18. Zhang J, An C, Dai J, et al. Joint vessel segmentation and deformable registration on multi-modal retinal images based on style transfer. *IEEE Int Conf Image Process*. 2019;839–843.
 19. Wang Y, Zhang J, An C, et al. A segmentation based robust deep learning framework for multi-modal retinal image registration. *Proc. IEEE Int. Conf. Acoust. Speech Signal Process*; 2020.
 20. Gatys LA, Ecker AS, Bethge M. A neural algorithm of artistic style. *CoRR*. 2015;1508; abstract 06576.
 21. Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. In *Computer Vision. ECCV*. 2016;694–711. Available at: https://link.springer.com/chapter/10.1007/978-3-319-46475-6_43#citeas.
 22. Staal J, Abràmoff MD, Niemeijer M, Viergever MA, Van Ginneken B. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans Med Imag*. 2004;23(4):501–509.
 23. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention. MICCAI*; 2015;234–241.
 24. DeTone D, Malisiewicz T, Rabinovich A. Superpoint: self-supervised interest point detection and description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 224–236.
 25. Yi KM, Trulls E, Ono Y, Lepetit V, Salzmann M, Fua P. Learning to find good correspondences. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 2666–2674.
 26. Rohm M, Tresp V, Muller M, et al. Predicting visual acuity by using machine learning in patients treated for neovascular age-related macular degeneration. *Ophthalmology*. 2018;125(7):1028–1036.
 27. US Food and Drug Administration (FDA). FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. *Case Med Res*. 2018, <https://doi.org/10.31525/fda2-ucm604357.htm>.
 28. Ying G-S, Huang J, Maguire MG, et al. Baseline predictors for one-year visual outcomes with ranibizumab or bevacizumab for neovascular age-related macular degeneration. *Ophthalmology*. 2013;120:122–129.

29. Ortega M, Gonzalez M, Rouco J, Barreira N, Carreira MJ. Retinal verification using a feature points based biometric pattern. *EURASIP Journal on Advances in Signal Processing*. 2009;13:2009. Article ID 235746.
30. Sauer L, Andersen KM, Li B, Gensure RH, Hammer M, Bernstein PS. Fluorescence lifetime imaging ophthalmoscopy (FLIO) of macular pigment. *Invest Ophthalmol Vis Sci*, <https://doi.org/10.1167/iovs.18-23886>.
31. Holz FG, Steinberg JS, Göbel A, Fleckenstein M, Schmitz-Valckenberg S. Fundus autofluorescence imaging in dry AMD: 2014 Jules Gonin lecture of the Retina Research Foundation. *Graefes Arch Clin Exp Ophthalmol*. 2015;253(1):7–16.
32. Ly A, Nivison-Smith L, Assaad N, Kalloniatis M. Infrared reflectance imaging in age-related macular degeneration. *Ophthalmic Physiol Opt*. 2016;36(3):303–316.
33. Greenstein VC, Ari D, Schuman J, et al. Near-infrared autofluorescence: its relationship to short-wavelength autofluorescence and optical coherence tomography in recessive Stargardt disease. *Invest Ophthalmol Vis Sci*. 2015;56(5):3226–3234.
34. Muftuoglu IK, Gaber R, Bartsch DU, Meshi A, Goldbaum M, Freeman WR. Comparison of conventional color fundus photography and multi-color imaging in choroidal or retinal lesions. *Graefes Arch Clin Exp Ophthalmol*. 2018;256(4):643–649.
35. Vallabh NA, Sahni JN, Parkes CK, Czanner G, Heimann H, Damato B. Near-infrared reflectance and autofluorescence imaging characteristics of choroidal nevi. *Eye (Lond)*. 2016;30(12):1593–1597.
36. Landa G, Rosen RB, Garcia PMT, Seiple WH. Combined three-dimensional spectral OCT/SLO topography and microperimetry: steps toward achieving functional spectral OCT/SLO. *Ophthalmic Res*. 2010;43(2):92–98.