Title

Efficacy and accuracy of artificial intelligence to overlay multimodal images from different optical instruments in patients with retinitis pigmentosa

Journal

Clinical and Experimental Ophthalmology, 51(5)

ISSN

1442-6404

Authors

Yassin, Shaden H

Wang, Yiqian

Freeman, William R

et al.

Publication Date

2023-07-01

DOI

10.1111/ceo.14234

Peer reviewed

HHS Public Access

Author manuscript

Clin Exp Ophthalmol. Author manuscript; available in PMC 2023 July 10.

Published in final edited form as:

Clin Exp Ophthalmol. 2023 July ; 51(5): 446–452. doi:10.1111/ceo.14234.

Efficacy and accuracy of artificial intelligence to overlay multimodal images from different optical instruments in patients with retinitis pigmentosa

Shaden H. Yassin, MD, PhD[1], **Yiqian Wang, PhD**[2], **William R. Freeman, MD**[1], **Anna Heinke, MD, PhD**[1], **Evan Walker, MSc**[1], **Truong Nguyen, PhD**[2], **Dirk-Uwe G. Bartsch, PhD, FARVO**[1], **Cheolhong An, MS**[2], **Shyamanga Borooah, FRCOphth, PhD**[1]

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

[2]Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, California, USA

Abstract

Background: Retinitis pigmentosa (RP) represents a group of progressive, genetically heterogenous blinding diseases. Recently, relationships between measures of retinal function and structure are needed to help identify outcome measures or biomarkers for clinical trials. The ability to align retinal multimodal images, taken on different platforms, will allow better understanding of this relationship. We investigate the efficacy of artificial intelligence (AI) in overlaying different multimodal retinal images in RP patients.

Methods: We overlayed infrared images from microperimetry on near-infrared images from scanning laser ophthalmoscope and spectral domain optical coherence tomography in RP patients using manual alignment and AI. The AI adopted a two-step framework and was trained on a separate dataset. Manual alignment was performed using in-house software that allowed labelling of six key points located at vessel bifurcations. Manual overlay was considered successful if the distance between same key points on the overlayed images was ≤1/2°.

Results: Fifty-seven eyes of 32 patients were included in the analysis. AI was significantly more accurate and successful in aligning images compared to manual alignment as confirmed by linear mixed-effects modelling ($p < 0.001$). A receiver operating characteristic analysis, used to compute the area under the curve of the AI (0.991) and manual (0.835) Dice coefficients in relation to their respective 'truth' values, found AI significantly more accurate in the overlay ($p < 0.001$).

Conclusion: AI was significantly more accurate than manual alignment in overlaying multimodal retinal imaging in RP patients and showed the potential to use AI algorithms for future multimodal clinical and research applications.

Correspondence: Shyamanga Borooah, Ophthalmology, Retinal division, Shiley Eye Institute, University of California, San Diego, 9415 Campus Point Dr, La Jolla, CA 92093, USA. sborooah@health.ucsd.edu.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

Keywords

artificial intelligence; microperimetry; retinitis pigmentosa; spectral domain optical coherence tomography; structure–function

1 | INTRODUCTION

Retinitis pigmentosa (RP) represents a diverse group of genetically heterogenous diseases that are progressive and can lead to blindness.[1] Recent developments in understanding RP, aim to establish relationships between measures of retinal function and structure, with the goals of understanding the correlation between retinal function at precise locations to help identify outcome measures or biomarkers for clinical trials.[2–5] Moreover, there is an increased opportunity to identify novel biomarkers with the increased use of multimodal imaging. Furthermore, computer learning may enable more sensitive quantification of key study end points compared to traditional measurements.[5,6]

Artificial intelligence (AI)-based algorithms are rapidly entering the healthcare field and previous studies have shown the remarkable potential to improve patient care.[7] Ophthalmology has been at the forefront of medical specialties adopting AI. However, these imaging techniques require the use of multiple devices with their own proprietary software, which is a barrier to the use of these imaging modalities in the clinic or research setting. Advances in AI now allow for the combination of multimodal imaging to be used as a co-localised database. AI has been previously used in a similar way to categorise or detect disease.[8–10] On the other hand, preliminary trials to overlay multimodal imaging, taken on different platforms, using Microsoft PowerPoint, was not successful (Figure 1).

To our knowledge, AI has not been used to combine structural images from scanning laser ophthalmoscope (SLO)/spectral domain optical coherence tomography (SD-OCT), over functional images from microperimetry. In this study, we aim to investigate the efficacy of AI to generate multi-instrument structural and functional imaging studies, by overlaying images from a cohort of patients diagnosed with RP.

2 | METHODS

This retrospective cross-sectional study used imaging of patients with RP seen in the inherited retinal dystrophy (IRD) clinic at the University of California San Diego (UCSD) from July 2021 to April 2022. The institutional review board approved this study for a waiver of informed consent which was granted due to the retrospective nature of this work. The study was conducted according to the principles of the Helsinki Declaration. RP patients were identified from the IRD database at UCSD, for image processing patient's identification information was redacted. Patients were included if they had a history of RP as confirmed by a retinal attending. After reviewing the images, the findings were confirmed by a retinal fellow. Eyes were selected if they had imaging including MAIA microperimetry (S-MAIA, Centervue, Italy), and SLO/SD-OCT (Spectralis HRA + OCT, Heidelberg Engineering, Heidelberg, Germany) performed on the same day and were of

good quality with vessels clearly seen on infrared (IR) images from microperimetry. Those not deemed of high enough quality by an experienced ophthalmologist were excluded.

IR images (850 nm) from microperimetry were overlayed on near infra-red (NIR) images (815 nm) of SLO/SD-OCT scans using manual alignment and AI. The AI adopted a two-step framework (coarse-to-fine) (Figure 2) using deep learning neural networks for the alignment of images and was initially trained using a separate dataset with 530 image pairs for training, 90 pairs for validation, and 253 pairs for testing. This training dataset contained a variety of pathologies including diabetic retinopathy, haemorrhages, and macular degeneration.[11] For co-registration, the coarse alignment network output the perspective transformation matrix, and the fine alignment network output a pixel-level deformation map that supported nonrigid transformations. The coarse network included three different sub-networks for segmentation, feature extraction, and outlier rejection. The network architecture had been described previously.[11] The vessel segmentation of the input IR images was obtained by the segmentation sub-network, as part of the coarse alignment network. The grayscale vessel segmentation map had values between [0, 1], characterising the probability of a pixel being vessel or not. The network was trained to resemble a manually labelled segmentation image, as the style target, and preserving the vessel structure of the input image. The input images were not cropped before sending into the networks, as the AI method was robust to field-of-view differences.

A software was developed to allow the authors to perform manual alignment by manually labelling six pairs of key points. These key points were located at vessel bifurcations. For co-registration, the manual alignment method used perspective transformation. The nature of the perspective transformation used included translation, scaling, shearing, and perspective distortion and required only four pairs of corresponding key points to estimate eight degrees of freedom. Note that the estimated transformation, in the manual approach, could not guarantee zero error on all six matching key points, as six pairs of key points were overdetermined for the perspective transformation and the estimation algorithm optimised for the least square error. The software provided alpha-blending and colour channel overlay results to allow the grader to check back-and-forth to achieve the best registration quality. The manual labelling was performed by two experienced ophthalmologists (S. Y., A. H.).

A software was developed to compare between AI and manual methods, recorded as a result of unique image pairs multiplied by 2 as block replicates, overlying both techniques; manual and AI. Each image was divided into 25 squares (5 × 5). Each image was a checkerboard composed of alternating IR microperimetry and N-IR SLO/SD-OCT Images (Figure 2).

To compare the alignment between manual and AI, the pixel error was measured on the labelled key point pairs. The overlay was considered successful if the maximum distance between all matched key points on the overlayed images was ≤½°. The ½° size was used, as the microperimetry stimuli use a Goldman size III stimulus which is ~½° in size, this results in ~100 microns on the retina and thus would be highly accurate for microperimetry purposes. Besides the maximum distance, the Sørensen–Dice coefficient of the aligned images was also used, where a higher number indicated a greater ratio of overlapping vessels segmentation between the aligned IR images. The Dice coefficient would be 1 if both perfect

vessel segmentation and perfect alignment. The Dice coefficient used checked the alignment of the target versus the source vessel map and was defined as

$$\mathrm{DICE}(I_1, I_2) = \frac{2|I_1 \cap I_2|}{|I_1| + |I_2|}$$

I_1 and I_2 were binary images in this function. Since the vessel maps were in grayscale, a threshold at 0.5 was used to convert the images to binary. This methodology was previously described in the study by Zhang et al.[11] With the given vessel segmentation method, a better alignment matrix would result in higher Dice coefficient for each image pair. In this sense a ground truth alignment for evaluation was not required in determining the Dice coefficient.

We utilised linear mixed-effects models, as well as a receiver-operating characteristic (ROC) analysis, to compare between both methods using the statistical software R (Version 4.2.1). The area under the ROC curve (AUC) of the AI and manual Dice coefficients in relation to their respective 'truth' values was evaluated.

3 | RESULTS

Image pairs of 69 eyes (three patients had only one eye imaged) from 36 consecutive series of RP cases taken using both microperimetry and SLO/SD-OCT on the same day were first reviewed. Two image pairs were excluded because they were not taken on the same day, and 12 image pairs were excluded due to poor quality of the MAIA infra-red images. Of the total of 69 reviewed image pairs, 57 images pairs of RP eyes were overlayed using both manual alignment and AI.

In order to determine which method performed best at overlaying the images, we used linear mixed-effects models, with random intercepts to account for intercorrelated subject eyes, to generate mean and 95% confidence intervals for comparison between AI Dice coefficients (mean 0.52 (95% CI 0.46, 0.57)) versus the manual Dice coefficient (mean 0.34 (95% CI 0.30, 0.37)). A statistically significant mean difference of 0.18 (95% CI 0.14, 0.23) between the AI and manual Dice coefficients was found ($p < 0.001$). (Figure 3 represents two examples comparing the accuracy of the manual vs. the AI methods).

To compute the area under the ROC curve of the AI and manual Dice coefficients in relation to their respective 'truth' values (defined by error less than Goldmann III), a ROC analysis (Figure 4) was used. AI Dice coefficient AUC (95% CI) = 0.991 (0.957, 1.000), whereas the manual Dice coefficient AUC (95% CI) = 0.835 (0.691, 0.940). Thus, AI was found to be more accurate than the manual alignment method. Using a bootstrapped test for two ROC curves, a significant difference ($p < 0.001$) in AUC's was found between these two methods. Bootstrap resampling was performed on a patient level, and a sample size of 2000 bootstrap resamples was used for evaluation of difference in AUC's.

4 | DISCUSSION

In the current study we investigated the accuracy of AI compared to manual alignment to overlay IR images from MAIA microperimetry and N-IR images from SLO/SD-OCT

scans in patients with RP. The results demonstrate that AI was more accurate than manual alignment in overlaying images from different optical systems.

There is a relative lack of papers describing retinal image overlay using AI. The efficacy of AI to overlay retinal images taken by different imaging systems has previously been investigated by Cavichini et al.[10] In that study from our group, Cavichini et al. overlayed matched pairs of colour fundus (CF) photographs and infrared scanning laser ophthalmoscope (IR SLO) images using conventional alignment methods and AI. The AI overlay strategy used consisted of a joint vessel segmentation and a deformable registration model based on the convolutional neural network (CNN). Also, the AI was trained using previously published vessel segmentation. The conventional alignment method is a non-AI method that used mathematical warping algorithms to align vascular landmarks. They developed software where each image was divided into 25 squares (5 × 5). Each image was a checkerboard composed of alternating infrared and CF pictures and was graded twice. Each individual square was graded based on the alignment of the vessels. Following the comparison between both methods, AI achieved alignment matches in 94.4% cases versus 90.5% using the conventional method.[10]

Arikan et al. used vessel segmentation and automatic landmark detection to establish image registration, the process of aligning two or more images based on image appearances. They developed an algorithm allowing for tracing changes in retinal structure across different retinal imaging modalities.[12] They performed automatic multi-modal retinal 2D/3D image registration using U-Net for vessel segmentation and mask regional CNNs (R-CNN) for detection of vessel landmarks (bifurcations, branches and crossover). They then created two lists containing the coordinates of the found markers for the moving and the fixed images, followed by applying a dilation filter on the skeleton image to facilitate Dice comparison. For initial registration, they selected combinations of fixed and moving images with highest Dice values. They applied their method on different combinations of modalities including fundus autofluorescence (FAF) to scanning laser ophthalmoscopy (SLO)/OCT, fluorescein angiogram (FA) to OCT-angiography (OCT-A), and indocyanine green angiography (ICGA) to FA and compared their algorithm with intensity based multimodal affine registrations from elastix. They achieved an average error rate of 13.12% versus 30.8% using the reference method, which confirmed that multi-modal registration of retinal images using vessel segmentation and landmark detection was more accurate than reference methods.[12] Lee et al. introduced a non-human dependent method for multimodal retinal image registration that is based on CNN.[13] They developed a deep step patch pattern (DeepSPa) framework, where they used intersection points between edges being formed by vascular junctions in the retinal images. Following the extraction of pattern patches surrounding the intersection points, they sorted the pattern patches into classes according to their pixel patch patterns. CNN was used to learn the patches to be used for matching. They then applied their approach to FA, OCT fundus and CF images. The comparative tests were demonstrated on three multimodal retinal image datasets; CF-FA (mild-to-moderate retinal diseases), CF-FA (severe retinal diseases), and CF-OCT (speckle noise and low resolution). Their results indicate that DeepSPa achieves higher registration accuracy than the other six existing algorithms in their experiments. It has also consistently attained above 80% registration success rates on all the three datasets, confirming its accuracy.[13]

To our knowledge, the current study is the first to investigate the efficacy of AI to overlay structural over functional retinal images in patients with RP compared with manual alignment. Our results show a high accuracy of AI in multimodal retinal image registration using functional testing techniques (microperimetry) and structural SLO/SD-OCT testing. This constitutes an important step towards developing efficient algorithms for future clinical implications. Using such precision overlays, it will be possible to correlate and monitor functional deficits as evidenced by microperimetric retinal sensitivity in small areas with structural changes in retinal degenerations.

FUNDING INFORMATION

Shyamanga Borooah is funded by a Foundation Fighting Blindness Career development award, and a Nixon Visions Foundation award. Truong Nguyen is funded by the National Eye Institute grant R01EY033847. This work was funded by NIH grants R01EY033847. William R. Freeman and Truong Nguyen are funded by R01EY 016323 and core grant P30EY022589.

REFERENCES

1. Duncan JL, Pierce EA, Laster AM, et al. Inherited retinal degenerations: current landscape and knowledge gaps. Transl Vis Sci Technol. 2018;7(4):6.
2. Lad EM, Duncan JL, Liang W, et al. Baseline microperimetry and OCT in the RUSH2A study: structure-function association and correlation with disease severity. Am J Ophthalmol. 2022;244:98–116. [PubMed: 36007554]
3. Birch DG, Cheng P, Duncan JL, et al. The RUSH2A study: best-corrected visual acuity, full-field electroretinography amplitudes, and full-field stimulus thresholds at baseline. Transl Vis Sci Technol. 2020;9(11):9.
4. Birch DG, Samarakoon L, Melia M, et al. The RUSH2A study: dark-adapted visual fields in patients with retinal degeneration associated with biallelic variants in the *USH2A* gene. Invest Opthalmol Vis Sci. 2022;63(3):17.
5. Wang YZ, Birch DG. Performance of deep learning models in automatic measurement of ellipsoid zone area on baseline optical coherence tomography (OCT) images from the rate of progression of USH2A-related retinal degeneration (RUSH2A) study. Front Med. 2022;5(9):932498.
6. Lee CS, Lee AY. How artificial intelligence can transform randomized controlled trials. Transl Vis Sci Technol. 2020;9(2):9.
7. Alugubelli R. Exploratory study of artificial intelligence in healthcare. Int J Innov Eng Res Technol. 2016;3(1):11.
8. Arcadu F, Benmansour F, Maunz A, et al. Deep learning predicts OCT measures of diabetic macular thickening from color fundus photographs. Invest Opthalmol Vis Sci. 2019;60(4): 852–857.
9. Mariottoni EB, Datta S, Dov D, et al. Artificial intelligence mapping of structure to function in glaucoma. Transl Vis Sci Technol. 2020;9(2):1–14.
10. Cavichini M, An C, Bartsch DUG, et al. Artificial intelligence for automated overlay of fundus camera and scanning laser ophthalmoscope images. Transl Vis Sci Technol. 2020;9(2):1–10.
11. Zhang J, Wang Y, Dai J, et al. Two-step registration on multimodal retinal images via deep neural networks. IEEE Trans Image Process. 2022;31:823–838. [PubMed: 34932479]
12. Arikan M, Sadeghipour A, Gerendas B, Told R, Schmidt-Erfurt U. Deep learning based multi-modal registration for retinal imaging. In: Suzuki K, Reyes M, Syeda-Mahmood T, et al., eds. Interpretability of Machine Intelligence in Medical Image Computing and Multimodal Learning for Clinical Decision Support [Internet]. Lecture Notes in Computer Science. Vol 11797. Springer International Publishing; 2019:75–82. doi:10.1007/978-3-030-33850-3_9
13. Lee J, Liu P, Cheng J, Fu H. A deep step pattern representation for multimodal retinal image registration. 2019 IEEE/CVF International Conference on Computer Vision (ICCV) [Internet]. IEEE; 2019:5076–5085 https://ieeexplore.ieee.org/document/9008309/

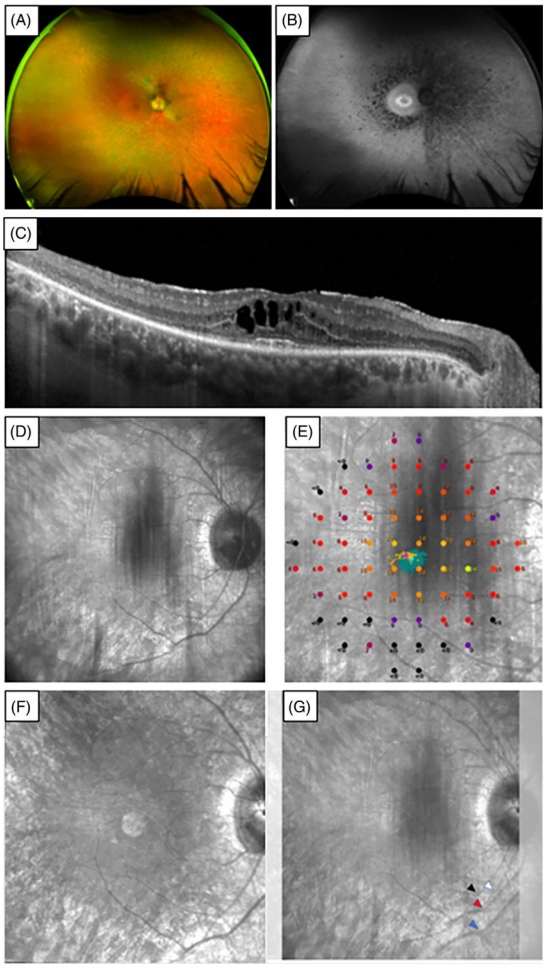

FIGURE 1.

Multimodal imaging of the right eye of a retinitis pigmentosa patient (A) wide field pseudocolor fundus photo (B) wide field fundus autofluorescence. (C) Horizontal spectral domain optical coherence tomography (SD-OCT) scan. (D) Infrared (IR) microperimetry image. (E) Microperimetry sensitivity scores. (F) N-IR of SD-OCT. (G) Overlay of microperimetry IR over SD-OCT IR showing discrepancies due to differences in sizing and orientation of images with markings showing corresponding points in retinal arteries (black and white arrowheads) and veins (blue and red arrowheads). The temporal, superior and inferior disc rims were well aligned.

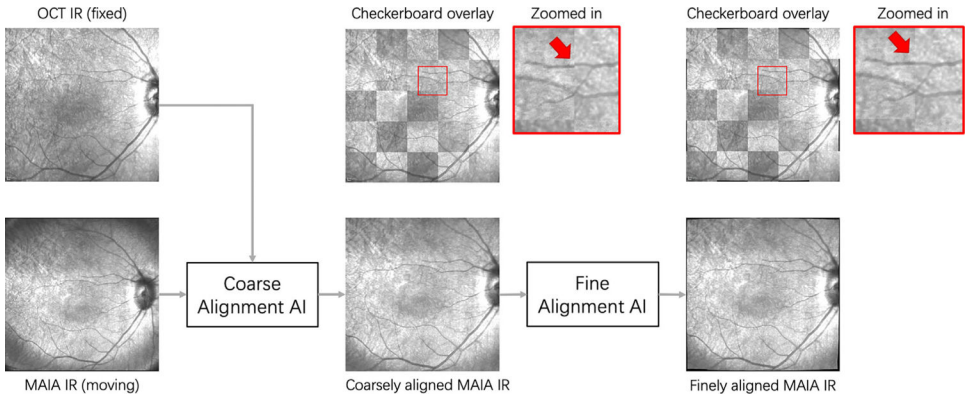

FIGURE 2.
The coarse-to-fine alignment artificial intelligence (AI) framework.

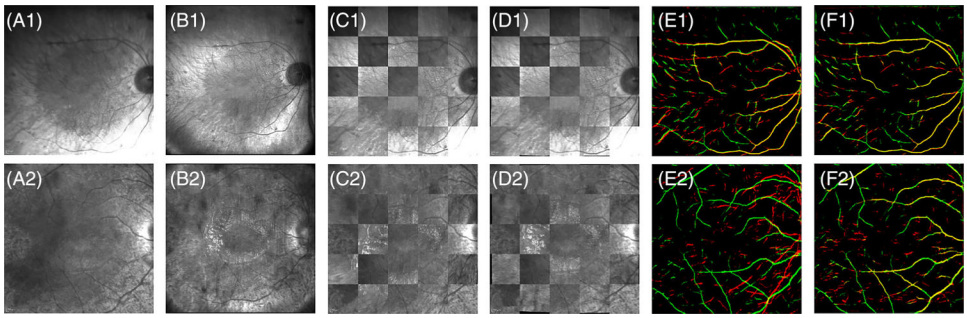

FIGURE 3.

Comparing the accuracy of manual alignment and artificial intelligence (AI). Row (1), and (2) show two examples; row (1) both AI and manual methods were successful in the alignment, but AI was more accurate, and in row (2) only AI was successful to overlay the images. Column (A) optical coherence tomography (OCT) IR image, (B) MAIA IR image, (C) checkerboard overlay by manual alignment, (D) checkerboard overlay by AI alignment, (E) vessel segmentation overlay by manual alignment (green: OCT, red: MAIA, yellow: overlap), (F) vessel segmentation overlay by AI alignment.

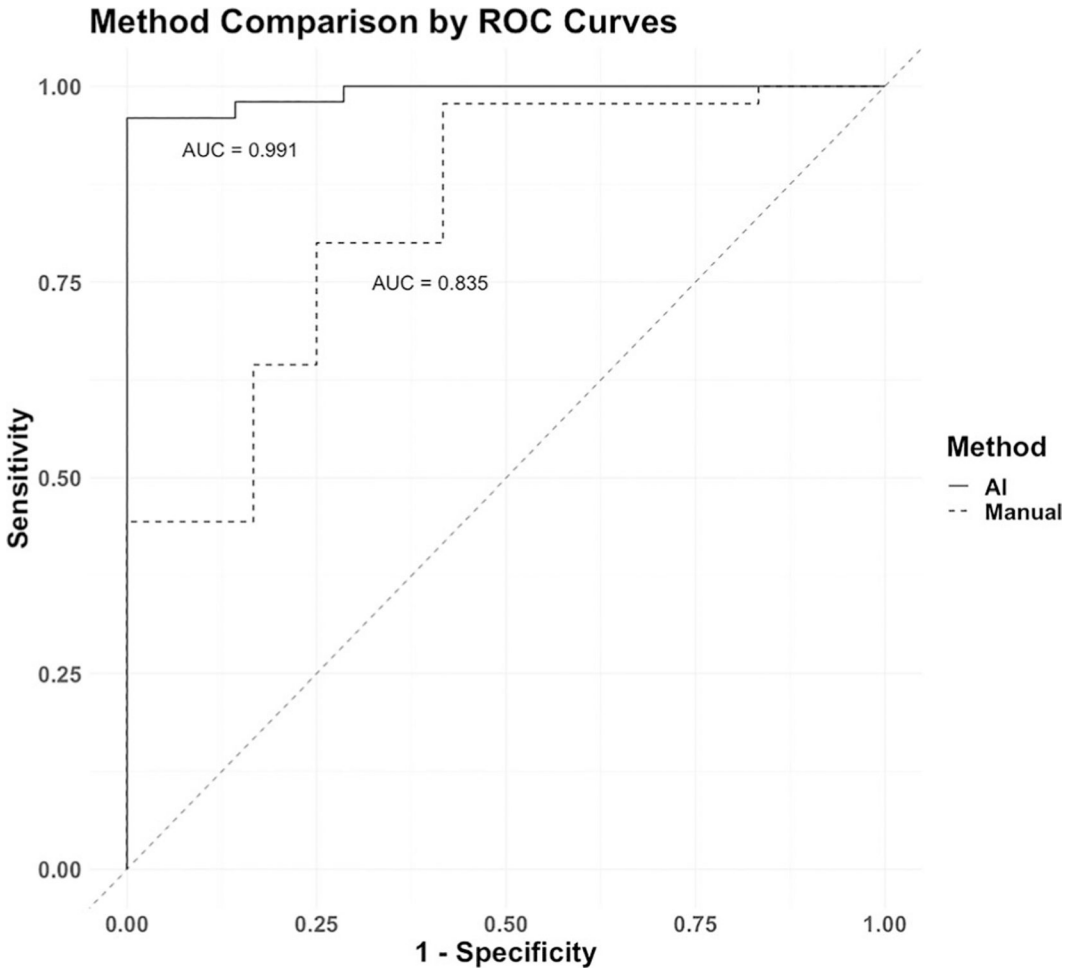

FIGURE 4.
Receiver-operating characteristic (ROC) analysis comparing area under the curve (AUC) of the artificial intelligence (AI) and manual Dice coefficients.