

USING AI TO IDENTIFY EXUDATIVE MACULAR DISEASE



A new tool for OCT interpretation was able to detect macular edema secondary to retinal pathology with high accuracy.

BY CLAUDIA P. ACOSTA, MD; MARIA A. PIEDRAHITA, MD; JUAN GONZALO SANCHEZ, MD; JULIANA MUÑOZ-ORTIZ, MD; JULIÁN MARTINEZ; JOSÉ ZEA; LAURA VELASQUEZ; JULIANA MESA; SEBASTIAN MENGES; AND CAROLINA SARDI, MD, MS

Diseases affecting the macula must be diagnosed quickly to establish immediate management and prevent further vision loss.^{1,2} AMD, diabetic macular edema (DME), and vascular occlusion are known to have an exudative component that generates fluid accumulation in the retina, and OCT is the imaging modality of choice to diagnose these retinal pathologies, as it is able to identify macular edema.^{3,4}

A tool to support OCT interpretation would potentially help avoid progression to severe exudative disease and allow retina specialists to identify patients who have intraretinal fluid and, thus, may benefit from a careful retina examination and/or require antiangiogenic therapy.³⁻⁶

In this article, we describe a convolutional neural network (CNN) we developed that may be able to identify patients with intraretinal fluid secondary to exudative disease based on the interpretation of OCT B-scans.

STUDY DESIGN

We conducted a retrospective secondary source analysis study with the goal of developing an AI model that would be able to distinguish between patients who have intraretinal fluid and macular edema and those who do not. The study involved first training a neural network to assess the selected OCT scans and then validating its ability for image discrimination.

Image Classification and Processing

Images of 158 eyes (108 patients) were included for training: 39 normal eyes, 24 with DME, 62 with wet AMD, and 33 with nonexudative retinal pathologies. Two retina specialists had examined the images to ascertain whether any of them contained biological markers indicative of intraretinal fluid secondary to macular exudative pathology. Images with artifacts or signal strength lower than 6/10 were excluded, and linear desegmentation was removed.

For training purposes, 110/158 of the images were used, while 28 were allocated to testing and 20 were used for validation. Data augmentation was performed to generate up to 2,000 images from the original 110 images to train the algorithm. This augmentation consisted of a series of random image transformations, such as rotations, brightness changes, shear transformations, zoom changes, and horizontal flips.

RESULTS

Among the 28 images used for testing, the algorithm correctly classified 27 as true positives and misclassified one (Figure 1). The algorithm presented an odds ratio of 14.81 ($P < .001$), which indicates a significant association between the predicted and actual labels, suggesting the algorithm's performance was not due to chance.

The evaluation metrics obtained for the algorithm had an area under the curve (AUC) of 0.965, precision of 0.979,

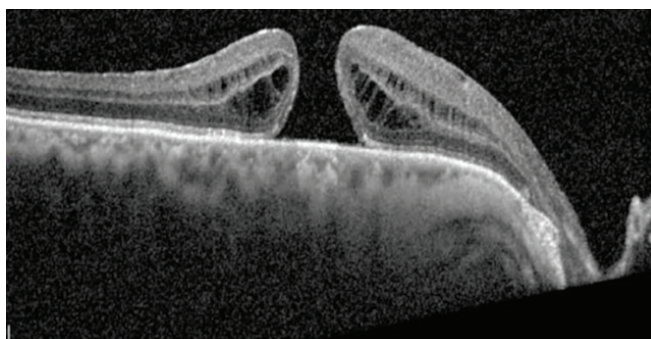


Figure 1. This OCT was the image that the algorithm misclassified. It shows a total thickness macular hole with intraretinal cysts on the edge of the hole, which does not correspond to a pathology involving macular edema.

accuracy of 0.964, recall/sensitivity of 0.900, F1 score of 0.934, and specificity of 0.900 (Figure 2). These metrics resulted in a global score of 94.0. The accuracy, F1 score precision, recall/sensitivity, specificity, and AUC values were calculated using the 28 test images.

To further assess the algorithm's reliability, a nonparametric confidence interval calculation was performed over a different set of 20 randomly selected images from the dataset to replicate the testing process. The 95% confidence intervals calculated for accuracy, F1 score, precision, recall/sensitivity, specificity, and AUC were [0.8, 1.0], [0.0, 1.0], [0.0, 1.0], [0.0, 1.0], [0.8, 1.0], and [0.5, 1.0], respectively. These results indicate a high level of confidence that the evaluation metrics obtained were accurate and representative of the model's performance on the population of images.

ALGORITHMS IN THE LITERATURE

Many investigators are working on developing new machine-learning algorithms for OCT interpretation. Grassman et al developed an algorithm to classify AMD using the AREDs scale that included 86,770 images.⁷ They found a sensitivity and specificity of 0.943 and 0.842 for no AMD, 0.822 and 0.971 for intermediate AMD, and 1.0 and 0.965 for late-stage AMD, respectively.

Heo et al developed an algorithm for differentiating dry and wet AMD that included 399 images.⁸ Using five-fold cross-validation, the sensitivity was 0.9252 for dry AMD and 0.968 for wet AMD with a specificity of 0.917 for dry AMD and 1.0 for wet AMD.

In another study by Habra et al,⁵ an AI algorithm was trained to detect intraretinal and subretinal fluid in patients with wet AMD and DME using 15,482 OCT images from 1,185 patients. For the detection of any fluid, the sensitivity was between 0.84 and 0.95 with a specificity between 0.88 and 0.95 for each disease.

Unlike these algorithms, our CNN can determine whether a patient has macular edema in addition to intraretinal fluid. This training was accomplished using a smaller number of images due to the learning process of our algorithm.

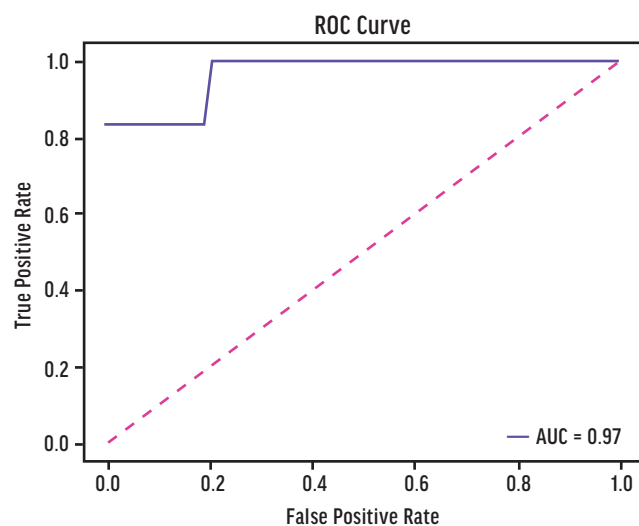


Figure 2. This chart represents the capability of the algorithm to classify the images correctly. The purple dotted line represents a random classifier with a 50% chance of success. The blue line represents the actual model performance. Characteristics such as a high AUC of 0.97 and the blue curve being closer to the top left corner of the graph indicate the algorithm's high accuracy in image classification.

STUDY LIMITATIONS

In training our algorithm, images with low signal-to-noise ratio or artifacts were excluded, which limits its use to high-quality images. Although the creation and training of the algorithm was successful and had a specificity and sensitivity of nearly 100%, it is essential to remember that these are tools to help the trained clinician make a diagnosis. The answers AI provides merely serve as a starting point, and the clinician must make the final determination on whether the patient has exudative maculopathy.

REAL-TIME EVALUATION IS NEEDED

This study validated our algorithm's ability to detect exudation in macular pathologies. Use of this algorithm may expedite clinical decision making and lead to earlier treatment. Prospective studies are needed to evaluate the performance of this neural network in a real-time environment. ■

1. Del Amo EM, Rimpelä AK, Heikkinen E, Kari OK, et al. Pharmacokinetic aspects of retinal drug delivery. *Prog Retin Eye Res.* 2017;57:134-185.
2. Ammar MJ, Hsu J, Chiang A, Ho AC, Regillo CD. Age-related macular degeneration therapy: a review. *Curr Opin Ophthalmol.* 2020;31(3):215-221.
3. Michl M, Fabianska M, Seeböck P, et al. Automated quantification of macular fluid in retinal diseases and their response to anti-VEGF therapy. *Br J Ophthalmol.* 2022;106(1):113-120.
4. Keenan TDL, Chakravarthy U, Loewenstein A, Chew EY, Schmidt-Erfurth U. Automated quantitative assessment of retinal fluid volumes as important biomarkers in neovascular age-related macular degeneration. *Am J Ophthalmol.* 2021;224:267-281.
5. Habra O, Gallardo M, Meyer Zu Westram T, et al. Evaluation of an artificial intelligence-based detector of sub- and intraretinal fluid on a large set of optical coherence tomography volumes in age-related macular degeneration and diabetic macular edema. *Ophthalmologica.* 2022;245(6):516-527.
6. Martin DF, Maguire MG, Fine SL, et al. Ranibizumab and bevacizumab for treatment of neovascular age-related macular degeneration: two-year results. *Ophthalmology.* 2012;119(7):1388-1398.
7. Grassmann F, Mengelkamp J, Brandl C, et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology.* 2018;125(9):1410-1420.
8. Heo TY, Kim KM, Min HK, et al. Development of a deep-learning-based artificial intelligence tool for differential diagnosis between dry and neovascular age-related macular degeneration. *Diagnostics.* 2020;10(5):261.

CLAUDIA P. ACOSTA, MD

- Retina Specialist, CEO, Instituto Nacional de Investigación en Oftalmología, Medellín, Colombia
- cacosta@inio.com.co
- Financial disclosure: None

JULIANA MUÑOZ-ORTIZ, MD

- Ophthalmologist, Escuela Superior de Oftalmología, Instituto Barraquer de América, Bogotá, Colombia
- Financial disclosure: None

JULIÁN MARTINEZ

- ML Engineer, Arkangel AI, Medellín, Colombia
- julian.martinez@arkangel.ai
- Financial disclosure: None

SEBASTIÁN MENGES

- Medical Student, Facultad de Medicina, Universidad CES, Medellín, Colombia
- menges.sebastian@uces.edu.co
- Financial disclosure: None

JULIANA MESA

- Medical Student, Facultad de Medicina, Universidad CES, Medellín, Colombia
- mesaac.juliana@uces.edu.co
- Financial disclosure: None

MARIA A. PIEDRAHITA, MD

- Ophthalmology Resident, Facultad de Medicina, Universidad El Bosque, Bogotá D.C., Colombia
- mapipiedrahita@gmail.com
- Financial disclosure: None

JUAN GONZALO SANCHEZ, MD

- Retina Specialist, Oftalmología de Precisión, Instituto Nacional de Investigación en Oftalmología, Medellín, Colombia
- juangosa@inio.com.co
- Financial disclosure: None

CAROLINA SARDI, MD, MS

- Retina Specialist, Epidemiologist, Instituto Nacional de Investigación en Oftalmología, Medellín, Colombia
- csardi@inio.com.co
- Financial disclosure: None

LAURA VELÁSQUEZ

- ML Engineer, Arkangel AI, Medellín, Colombia
- laura@arkangel.ai
- Financial disclosure: None

JOSÉ ZEA

- ML Engineer Arkangel AI, Medellín, Colombia
- jose@arkangel.ai
- Financial disclosure: None