AMERICAN ACADEMY OF OPHTHALMOLOGY®

Autonomous Screening for Diabetic Macular Edema Using Deep Learning Processing of Retinal Images

Idan Bressler,[1] *Rachelle Aviv*,[1] *Danny Margalit*,[1] *Gal Yaakov Cohen, MD*,[2,3] *Tsontcho Ianchulev, MD, MPH*,[1,4] *Shravan V. Savant, MD*,[5,6] *David J. Ramsey, MD, PhD*,[5,6] *Zack Dvey-Aharon, PhD*[1]

Objective: To develop and validate a deep learning model for diabetic macular edema (DME) detection using color fundus imaging, which is applicable in a diverse, multidevice clinical setting.

Design: Evaluation of diagnostic test or technology.

Subjects: A deep learning model was trained for DME detection using the EyePACS dataset, consisting of 32 049 images from 15 892 patients. The average age was 55.02%, and 51% of the patients were women.

Methods: Data were randomly assigned, by participant, into development (n = 14 246) and validation (n = 1583) sets. Analysis was conducted on the single image, eye, and patient levels. Model performance was evaluated using sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC). Independent validation was further performed on the Indian Diabetic Retinopathy Image Dataset, as well as on new data.

Main Outcome Measures: Sensitivity, specificity, and AUC.

Results: At the image level, a sensitivity of 0.889 (95% confidence interval [CI]: 0.878, 0.900), a specificity of 0.889 (95% CI: 0.877, 0.900), and an AUC of 0.954 (95% CI: 0.949, 0.959) were achieved. At the eye level, a sensitivity of 0.905 (95% CI: 0.890, 0.920), a specificity of 0.902 (95% CI: 0.890, 0.913), and an AUC of 0.964 (95% CI: 0.958, 0.969) were achieved. At the patient level, a sensitivity of 0.900 (95% CI: 0.879, 0.917), a specificity of 0.900 (95% CI: 0.883, 0.911), and an AUC of 0.962 (95% CI: 0.955, 0.968) were achieved.

Conclusions: Diabetic macular edema can be detected from color fundus imaging with high performance on all analysis metrics. Automatic DME detection may simplify screening, leading to more encompassing screening for diabetic patients. Further prospective studies are necessary.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article.

Supplemental material available at www.ophthalmologyscience.org.

Diabetic macular edema (DME) is a complication of diabetes mellitus, closely associated with diabetic retinopathy (DR).[1] Diabetic macular edema is characterized by the accumulation of excess fluid in the extracellular space within the central macula,[2,3] and when untreated, ultimately leads to vision loss due to damage to the microvasculature and photoreceptors of the fovea, which is responsible for high-resolution visual acuity. Diabetic macular edema has a major impact on public health, affecting approximately 3.8% of the population,[4] with an incidence of >25% within 25 years of diagnosis of type 1 diabetes mellitus[5] and 25% within 9 years of diagnosis of type 2 diabetes mellitus.[6]

The ETDRS defined clinically significant macular edema (CSME) with specific anatomic criteria, which include retinal thickening or the presence of hard exudates within 500 μm of the fovea.[7] Interventions studied in the ETDRS treatment protocol, such as focal laser photocoagulation and intravitreal anti-VEGF therapy, have shown significant improvement in visual acuity and prognosis after treatment.[8–10] Therefore, early detection and intervention are crucial in providing the opportunity to achieve good patient outcomes.

It is recommended that all patients with diabetes are screened for DME every 1 to 2 years. Screening is conducted as a part of regular screening and management for DR and typically involves slit lamp examination of the dilated fundus or the use of color fundus photography.[11] While the diagnosis of DME is traditionally based on fundus photography and fluorescein angiography, OCT has increasingly been used to quantitate the extent of diabetes-related retinal thickening.[12] This modality is expensive, however, and generally available only at specialized eye clinics that can afford this technology; as such, it is not widely accessible for primary screening purposes. Thus, screening for DME continues to be performed by detecting classic patterns of findings in

https://doi.org/10.1016/j.xops.2025.100722
ISSN 2666-9145/25

color fundus imaging, namely exudates and associated macular thickening.[10]

Access and scalability are crucial elements of any population health screening program. The need for a specialized eye examination and sophisticated equipment has been a significant obstacle for streamlined DME screening, causing many patients to remain undiagnosed.[13,14] Machine learning and deep learning algorithms are ideal tools to address this limitation and empower an efficient and exponential screening process at the level of nonspecialized, primary care delivery.[15–17]

When considering machine learning solutions for widespread screening purposes, ensuring maximal generalizability of the model becomes crucial as the exact settings in which the model will be deployed are diverse and unknown. As such, models should be developed using a wide and diverse database and tested in multiple settings. The ability to demonstrate consistent results in external validation, i.e., datasets that are different from the training dataset in terms of locations and protocols, is crucial. This is in addition and compared with internal validation, which is done on a subset of the original dataset reserved for validation. If the original database contains biases that affect the model, internal validation may not expose them, emphasizing the need for external validation.

Methods for DR screening using deep learning algorithms that examine fundus images have shown promising results.[18,19] Furthermore, autonomous DR screening has received United States Food and Drug Administration approval,[20] opening the door to the implementation of further similar applications, such as DME screening. While OCT-based machine learning methods have shown good results in the detection of DME,[21–23] with some methods boasting almost 100% accuracy, the cost and limited availability of OCT technology limit its ability to be used as a screening tool on a large scale.

Meta-analysis of machine learning–based DME detection methods has suggested that limitations of previous studies include a lack of comparison between eyes of the same patient, as well as the widespread use of OCT in previous works, which may limit generalizability among the general population.[24] This study relies on fundus images, which are more common and easily attainable, as well as analysis on the patient level, which includes both eyes, attempting to close previous research gaps.

Other fundus imagery-based methods, which focus primarily on exudate detection, were previously developed,[25,26] as have been methods based on the entire CSME criteria.[27–31] This study will expand on these works by being validated on multiple datasets from multiple locations and using multiple definitions for CSME, showing good agreement with all of them. Furthermore, not only does the model developed exhibit excellent CSME detection, but also it does so at the patient level rather than at the eye level, making this work more clinically relevant from a functional screening standpoint, as patient-level results are the end goal for diagnostic model screening purposes.

Methods

Data

The main dataset utilized for training and validation was compiled and provided by EyePACS[32] and consisted of 45° angle fundus photography images and expert readings of said images. All images and data were deidentified according to the Health Insurance Portability and Accountability Act "Safe Harbor" before they were transferred to the researchers. This study was conducted in compliance with the tenets of the Declaration of Helsinki, and institutional review board exemption was obtained.

The EyePACS dataset contained up to 6 images per patient visit: 1 macula-centered image, 1 disk-centered image, and 1 centered image per eye (in which a central fixation image is fixated on the middle of a line connecting the foveola and the optic disc). Each eye underwent expert reading, which included but was not limited to detecting the presence of DME, grading the level of DR, and assessing the image quality. For example, if 2 images of an eye are of bad quality, while the third is of excellent quality, the eye will be deemed readable. It should be noted that this quality assessment was based on the overall readability of a given eye and does not guarantee that all images of an eye were the same quality. Images fully deemed unreadable (i.e., ungradable) by an expert were omitted from our analysis, as were disc-centered images, because these provide only a limited view of the macula. Images that were partially visible were deemed insufficient for full interpretation by EyePACS but were included in the dataset given that partial visibility does not preclude DME grading.

The method used to determine the presence of CSME in fundus images in the vast majority of the EyePACS dataset was Bresnick's criterion,[33] which is defined as the existence of hard exudates within 1 disc diameter from the center of the macula. A different method used in a minority of the dataset was the criterion presented in Litvin et al[34]; this defined CSME by dividing the macula into an 8-sector "pie" within 1 disc diameter of the center of the macula. If 3 sectors of the pie have hard exudates, or if hard exudates are present within 1 of 3 disc diameter from the center of the macula, it was graded to have likely CSME. Due to the latter method's higher level of stringency, all images diagnosed using the latter are definitionally positive according to the former. As such, and given that images diagnosed using the latter method constitute a minority, all images were considered as having comparable ground truths.

A comprehensive dataset for the purposes of training, validation, and testing was constructed from the EyePACS dataset, consisting of 32 049 images from 15 892 patients. Up to 2 images were taken for each eye from 2 different fields, one centered on the macula and another centered between the macula and the disc. The average age was 55.02 (10.21 standard deviation), and 51% of the patients were women (Table 1). Table S1 (available at www.ophthalmologyscience.org) shows the distribution of DME patients across DR levels; DME was only present for patients with more than mild DR, with approximately 49% of all images being DME positive. Additional statistics are given in Tables S2 and S3 (available at www.ophthalmologyscience.org).

An additional dataset used for validation was the Messidor-2 dataset,[35] also containing 45° angle fundus photography images and expert annotations for DME presence and DR level. The dataset consisted of 1748 macula-centered images from 874 examinations, of which 151 images (8.6%) were DME positive. Additional information is provided in Table S4 (available at www.ophthalmologyscience.org).

Table 1. Patient Numbers and Population Statistics for the EyePACS Dataset

Field	Images	Patients	Mean Age (SD)	Sex (% Female)	Ethnicity (Fraction)
Value	32 049	15 892	55.02 (10.21)	51	White = 0.55 (Hispanic = 0.93, non-Hispanic = 0.07); ethnicity not specified = 0.13 African Descent = 0.11 Indian subcontinent origin = 0.10 Asian = 0.03 Other = 0.08

SD = standard deviation.

External validation was further performed on the Indian Diabetic Retinopathy Image Dataset, as well as a dataset comprised of patient images from Lahey Hospital and Medical Center.

Preprocessing

Image preprocessing was performed in 2 steps. First, the image background was cut along the convex hull, which contains the circular border between the image and the background. Figure 1 shows an example result of this process. Secondly, each image was resized to 512 × 512 pixels.

Quality Assessment

A heuristic tool for image quality assessment was developed. The tool gives a quality score for an image using an aggregation of the visibility score from multiple areas within the fundus image. Figure 2 demonstrates a few examples of images and their respective scores, showing the correlation between score and visual image quality.

Model Training

The model used was a convolutional neural network (for full architecture, see Supplementary Appendix 5, available at www.ophthalmologyscience.org) with an Adam optimizer and a learning rate of 0.003. The model was trained on 2 GTX 2080 Ti graphic cards. The hyperparameters for the model training were chosen beforehand and not changed to prevent overfitting.

Figure 1. Example of image cropping; blue lines represent the cropping boundaries.

The data were divided into training, validation, and test datasets consisting of 80%, 10%, and 10% of the data, respectively. Customary morphologic and color augmentation was applied during training. Training was done such that images from a patient are designated to one of the sets exclusively to avoid data leakage.

A binary classification neural network was trained. Following previous works,[36–38] the model architecture was automatically fitted to best balance the model performance versus model complexity tradeoff, ending with an empirically found best-performing model of ~28 million parameters and a size of ~110 MB. Hyperparameter tuning was done using the validation set. All metrics described here were assessed on the training set.

After evaluation of the performance of the model on images of different quality, a total of 1978 images with the worst quality were filtered out (approximately 6.2% of the data).

Statistical Analysis

The metrics used for assessment were accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). For each metric, the bias-corrected and accelerated bootstrap method[39] was used to produce a 95% confidence interval.

Analysis Levels

Diabetic macular edema detection was done on 3 different levels. The first level assessed detection at the level of each individual image, which was the basic task for which the model was trained. The second level was detection for each eye, using both macula-centered and mid-disc/macula-centered image fields for a given eye. This method is akin to fundus-photo-based detection of DME performed by a human expert. In this approach, the best image for each eye (in terms of image quality as assessed by the image quality tool) was selected for analysis. The third level was the patient level. For clinical purposes, detection of DME in 1 eye is sufficient for referral to further checks; as such, the "worst of 2 eyes" approach was used.

External Validation

The module was further validated by 2 external teams. The first external validation set consists of clinical data collected from 50 patients with DR at the Lahey Hospital and Medical Center. Of these patients, 19 had DME confirmed by OCT and 31 did not. One macula-centered image was selected from each eye for analysis, and the evaluation was done by D.R. based on Bresnick's method. Diabetic macular edema detection analysis was then performed on all images using the proposed model. Performance was judged using the metrics mentioned above.

The assessment for the second external validation was conducted by G.C. using the Indian Diabetic Retinopathy Image

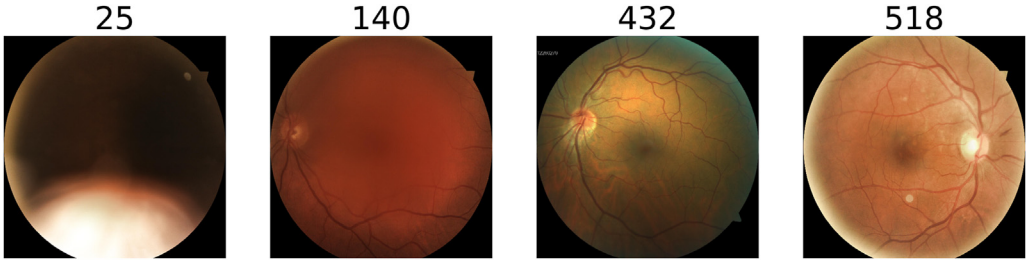

Figure 2. Example images and their accompanying image quality scores, ordered from worst quality (left) to best quality (right).

Dataset training dataset,[40] which is comprised of 400 macula-centered images. Diabetic macular edema and CSME detection were performed based on the method presented in Wong et al,[41] in which DME is defined via the existence of hard exudates 1 disc diameter from the macula, and CSME is defined via the existence of hard exudates 500 μm from the macula. The data were first annotated into 3 categories: DME positive (235), DME negative (142), and unreadable (23), and DME detection analysis was then performed on all readable images using the proposed model. Comparison was done using the metrics mentioned previously.

Model Visualization

The gradient-weighted class activation mapping method[42] was used to visualize model areas of interest.

Stability Analysis

In order to check the model's stability, a set of individual transformations was done on the input images, and the model's results for these transformed images were compared with the results of the original image.

The transformations were as follows:

A. Contrast change, by changing the brightness of the image by a factor of 0.05.
B. Rotation by 5 degrees
C. Scaling by zooming in by a factor of 0.15
D. Translation by a factor of 0.05 on both axes
E. Gaussian blur with a kernel size of 5 × 5.

An example of these transformations may be found in Figure 3.

Comparison was done by calculating the percent of changed classification results out of the test dataset for each transformation.

Results

EyePACS Dataset

The results for the different analysis methods are as follows (Table 2) (confidence intervals set to 95% in parentheses): on the image level, an AUC of 0.954 (0.949, 0.959) was achieved. On the eye level, an AUC of 0.964 (0.958, 0.969) was achieved. On the patient level, an AUC of 0.962 (0.955, 0.968) was achieved.

The results for each DR level for which DME is present are displayed in Table S6 (available at www.ophthalmologyscience.org), showing comparable results across all DR levels. The model achieved 0.958 AUC (0.952, 0.964) for DR level 2, 0.935 AUC (0.923, 0.945) for DR level 3, 0.940 AUC (0.926, 0.952) for DR level 4, and 0.954 AUC (0.919, 0.975) for ungradable DR level. Diabetic retinopathy grades 0 and 1 did not have any DME-positive examples; thus, most metrics are not defined for these grades; the model achieved an accuracy of 0.981 and 0.876 (confidence interval not defined), respectively.

Table S7 (available at www.ophthalmologyscience.org) shows the results for images that passed and did not pass the quality filter (high and low quality, respectively), showing significant differences between the populations. The results for images that were filtered out were 0.671 sensitivity (0.599, 0.737), 0.843 specificity (0.790, 0.886), and 0.853 AUC (0.811, 0.887). Results for images that passed the quality filter were 0.902 sensitivity (0.892, 0.912), 0.883 specificity (0.871, 0.893), and 0.956 AUC (0.952, 0.961) for images that passed the filter. The filter allowed for a reading on the patient level of 98% of the patient cohort.

Messidor-2 Dataset

Messidor-2, as a common benchmark, was used as an additional validation dataset. Messidor-2 contained readings for the image and patient levels, containing 1 image per eye. The results on this dataset were an AUC of 0.971 (0.955−0.982), a sensitivity of 0.875 (0.811−0.922), and a specificity of 0.954 (0.939−0.967), surpassing previous works (Table 3). On the patient level, an AUC of 0.964 (0.936, 0.979), sensitivity of 0.897 (0.820, 0.947), and specificity of 0.932 (0.905, 0.953) were achieved.

External Validation Datasets

When tested on the first validation dataset of 100 images, the model achieved 0.880 accuracy (0.757, 0.955), 0.789 sensitivity (0.544, 0.939), and 0.935 specificity (0.786, 0.992).

When tested on the second, larger validation dataset of 400 images, 23 were labeled as unreadable. Performance of the model on the remaining 377 images demonstrated 0.854 (0.812, 0.883) accuracy, 0.851 (0.802, 0.893) sensitivity, 0.859 (0.794, 0.909) specificity, and 0.931 (0.900, 0.953) AUC. For further information on the validation sets used, see Table S8 (available at www.ophthalmologyscience.org).

Model visualization (Fig 4) shows the model's focus areas using gradient-weighted class activation mapping, showing focus around the macula for both healthy and DME-positive images, as well as a focus on exudates in DME-positive images. Table S9 (available at

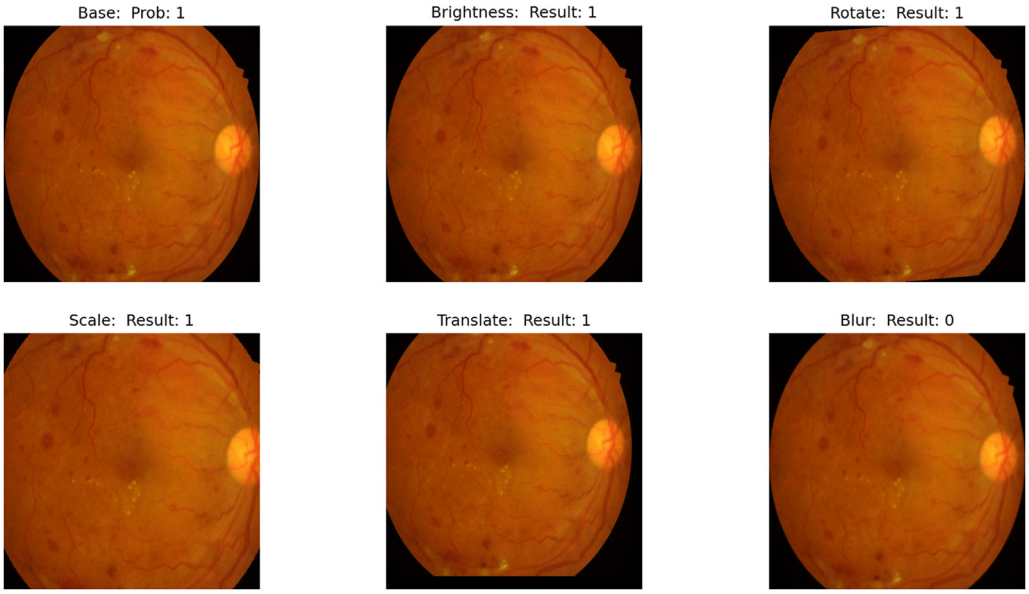

Figure 3. Example of transformations on the right eye from a DME-positive patient. Showing (clockwise): the original image, brightness shift, rotation, Gaussian blur, translation, and scale. The model result for each transformation is shown, where 1 is DME positive and 0 is DME negative. DME = diabetic macular edema.

www.ophthalmologyscience.org) shows the results for the stability analysis in terms of result change percentages for each transformation.

Discussion

This work introduces a novel, proprietary autonomous system for the detection of DME from fundus images. This may shorten and simplify the screening processes and allow for wider screening of DME. Given the short (within 3 months) recommended referral time after DME detection,[11] and the potential threat to patients' vision if left untreated, the widespread use of autonomous screening has the potential to be of clinical importance.

The need to screen for DME independently from autonomous DR screening stems from 3 main factors. Firstly, the recommended referral time for DME is shorter than that of most DR cases without DME.[11] Secondly, the treatment regime for DME differs from that of DR without DME,[43] emphasizing the importance of distinguishing DME cases from DR cases. Third, the effect of DME is usually more visually significant than DR and has a higher risk of causing irreversible vision changes.[44]

This work demonstrated good results on multiple validation datasets, from multiple locations, and with different definitions for CSME. This expands on previous works[45] by showing robustness in multiple different settings and across definitions used, demonstrating the applicability and general usability of this method. Our group has already incorporated a similar, United States Food and Drug Administration-cleared model for the screening of DR in primary care clinics using both handheld and tabletop cameras.[46,47] This model, if implemented in a similar manner, provides a practical screening solution.

While previous works have shown seemingly comparable results, these suffer from a few limitations that hinder the models' potential to act as widespread screening tools. Many works[28–31] have a significantly smaller dataset compared with that presented in this work, limiting diversity and generalizability. Homogeneity of data further limits generalizability. Sahlsten et al,[27] for instance, utilized a dataset based on a homogeneous population, both in terms of patients recruited and camera use: the work was done exclusively with a Canon camera and on a Finnish

Table 2. Results for the EyePACS Dataset across All 3 Analysis Levels, Given in Accuracy, Sensitivity, Specificity, and AUC with a 95% CI

	Accuracy (CI)	Sensitivity (CI)	Specificity (CI)	AUC (CI)
Image level	0.889 (0.881, 0.897)	0.889 (0.878, 0.900)	0.889 (0.877, 0.900)	0.954 (0.949, 0.959)
Eye level	0.903 (0.894, 0.912)	0.905 (0.890, 0.920)	0.902 (0.890, 0.913)	0.964 (0.958, 0.969)
Patient level	0.898 (0.886, 0.909)	0.900 (0.879, 0.917)	0.900 (0.883, 0.911)	0.962 (0.955, 0.968)

AUC = area under the receiver operating characteristic curve; CI = confidence interval.

Table 3. Comparison between the Proposed Method and Previous Works on the Messidor-2 Dataset, Given in Accuracy, Sensitivity, Specificity, and AUC with a 95% CI

	Accuracy (CI)	Sensitivity (CI)	Specificity (CI)	AUC (CI)
Sahlsten et al[27]	0.931 (0.915, 0.944)	0.69 (0.626, 0.750)	0.989 (0.980, 0.994)	0.932 (0.917, 0.946)
Li et al[28]	-	0.886 (0.881, 0.892)	0.908 (0.898, 0.912)	0.948 (0.943, 0.951)
Proposed, image level	0.943 (0.927, 0.955)	0.875 (0.811, 0.922)	0.954 (0.939, 0.967)	0.971 (0.955, 0.982)
Proposed, patient level	0.925 (0.898, 0.944)	0.897 (0.820, 0.947)	0.932 (0.905, 0.953)	0.964 (0.936, 0.979)

AUC = area under the receiver operating characteristic curve; CI = confidence interval.
Additionally, results on the patient level (not done in previous works) are given.

population, which is relatively less diverse than the United States-based population represented in this article.

Different sets in the training and validation sets use slightly different definitions for CSME; however, all use only exudates in macular proximity using fundus images. The comparable results between the training and external validation sets prove that the model is robust to these slight changes in definitions.

This work was validated on 4 validation sets that were diverse in terms of patient population and device (see Table S7, available at www.ophthalmologyscience.org), 3 of which were external, showing consistency in overall results and general consistency in terms of sensitivity–specificity balance. Balanced systems are important for screening purposes, as an imbalanced system will hurt either sensitivity or specificity beyond acceptable bounds, creating false positives or negatives during screening. Directly compared with other works that also utilize the Messidor set for external validation,[27,28] this work shows superior and significantly more balanced results.

Additionally, which further tailors this work to practical screening purposes, this work proposed analyzing DME on multiple levels, expanding on existing works that focused on the single image level, and shows higher efficacy on the image level as compared with previous studies. Additionally, it showed comparable results between the Messidor-2 dataset and the less curated (in terms of image quality) EyePACS dataset, demonstrating its robustness across different image qualities. The model can produce results for the vast majority of examined patients, further supporting the possible widespread capabilities and applications.

Analysis on the eye level, that is, analyzing a single eye with multiple images of the same eye, may be more accurate and representative of clinical practice than image-level analysis. When multiple fields of the same eye exist, experts label images based on the integration of present information. This may lead to the labeling of individual images being misleading, especially if differences in image quality exist. For instance, an eye that appears healthy from one angle, often due to low image quality, might have visible DME at another angle, leading to a seemingly

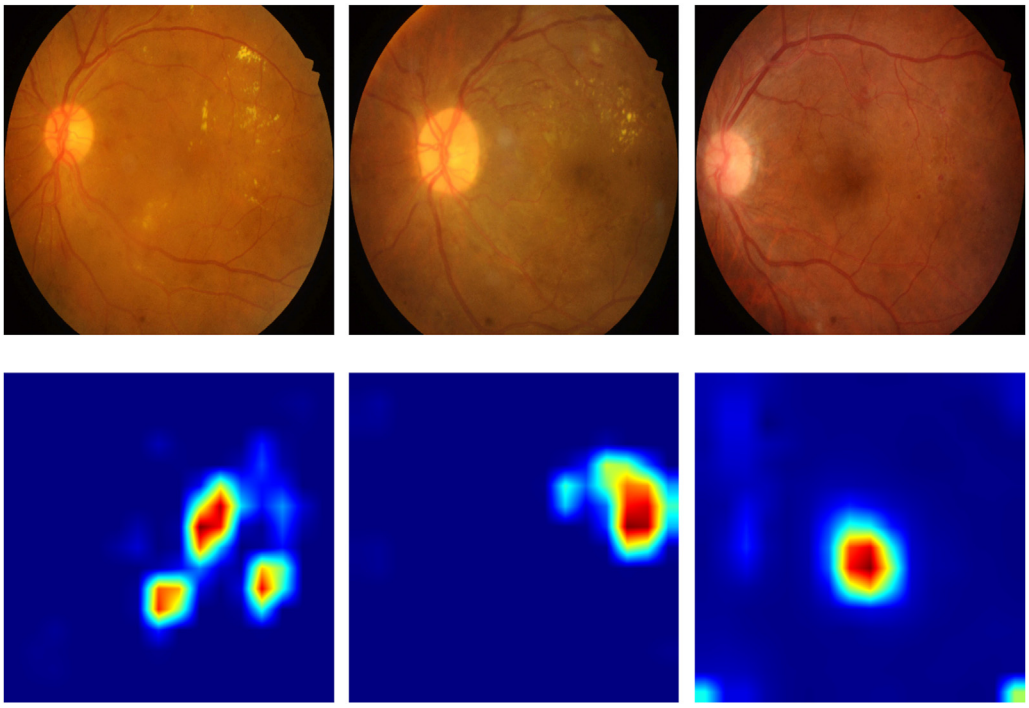

Figure 4. Grad-CAM heatmap results from 3 left eyes of 3 patients, with 22 confirmed DME patients (left and center) and 1 non-DME patient (right). DME = diabetic macular edema; Grad-CAM = gradient-weighted class activation mapping.

healthy image being positively labeled. The presented eye-level analysis tackles this issue by selecting the highest quality field from each eye.

The final model presented, which performs an analysis on the patient level, may be more clinically relevant than reporting findings at the single image or eye levels because the clinical criterion for referral is the existence of DME on the patient level. This method demonstrated high (~90%) sensitivity and specificity. Further analysis of the model shows good stability in response to image transformation, with Gaussian noise, which blurs image exudates, being the most problematic one; however, this is not a common artifact to find in retinal images.

Clinically significant macular edema with foveal involvement is also known as CSME with center involvement,[2,11] and the ability of graders to consistently detect this has been questioned.[48] Despite CSME with center involvement being more severe, all CSME cases are referable, and the detection of CSME remains common clinical practice and a referral marker, thus making widespread screening of CSME important. This article therefore focuses on CSME detection and not CSME with center involvement detection.

This work has a few limitations. Firstly, model training was performed on the single image level, thus hindering the training with the aforementioned image labeling problem. Secondly, the 2 methods used for ground truth may not be as accurate as a comprehensive eye examination using OCT in addition to the color fundus images. Additionally, a prospective study on noncurated data could cement the viability of this method for practical, everyday DME screening.

In conclusion, this article introduces a state-of-the-art method for DME detection and uniquely demonstrates clinical relevance by demonstrating results on multiple analysis levels. Furthermore, it demonstrates comparable and balanced results on multiple heterogeneous datasets, emphasizing the flexibility of the model.

Footnotes and Disclosures

Originally received: October 13, 2024.
Final revision: January 5, 2025.
Accepted: January 27, 2025.
Available online: January 31, 2025. Manuscript no. XOPS-D-24-00432.

[1] AEYE Health, Inc., New York, New York.
[2] The Goldschleger Eye Institute, Sheba Medical Center, Tel Hashomer, Israel.
[3] Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv, Israel.
[4] New York Eye and Ear, Mount Sinai Hospital, New York.
[5] Department of Ophthalmology, Lahey Hospital & Medical Center, Peabody, Massachusetts.
[6] Department of Ophthalmology, Tufts University School of Medicine, Boston, Massachusetts.

Disclosures:

All authors have completed and submitted the ICMJE disclosures form.

The authors made the following disclosures:

D.M.: Stocks and financial interest − AEYE Health, Inc. (COO and serves on AEYE Health's board)

G.Y.C.: Consultant − AEYE Health, Inc.

I.B.: Stocks − AEYE Health, Inc.

R.A.: Stocks − AEYE Health, Inc.

T.I.: Stocks and board member − AEYE Health, Inc.

S.V.S.: Financial interest − 2021 and 2022: dinners sponsored by Regeneron, Alcon, Sight Sciences, Ocular Therapeutix.

Z.D.-A.: Stocks − AEYE Health, Inc. (CEO and serves on AEYE Health's board).

This study was supported by AEYE Health, Inc., New York, USA.

Support for Open Access publication was provided by AEYE Health, Inc.

HUMAN SUBJECTS: No human subjects were included in this study. This study was conducted in compliance with the tenets of the Declaration of Helsinki, and institutional review board exemption was obtained.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Bressler, Margalit, Dvey-Aharon

Data collection: Savant, Ramsey

Analysis and interpretation: Bressler, Cohen, Savant, Ramsey, Dvey-Aharon

Obtained funding: N/A

Overall responsibility: Bressler, Aviv, Ianchulev, Dvey-Aharon

Abbreviations and Acronyms:
AUC = area under the receiver operating characteristic curve; **CSME** = clinically significant macular edema; **DME** = diabetic macular edema; **DR** = diabetic retinopathy.

Keywords:
Artificial intelligence, Diabetic macular edema, Deep learning, Fundus.

Correspondence:
Zack Dvey-Aharon, PhD, 1501 Broadway, New York City, NY 10036. E-mail: zack@aeyehealth.com.

References

1. Mohamed Q, Gillies MC, Wong TY. Management of diabetic retinopathy: a systematic review. *JAMA*. 2007;298:902−916.
2. Lang GE. Diabetic macular edema. *OPH*. 2012;227:21−29.
3. Bandello F, Parodi MB, Lanzetta P, et al. Diabetic macular edema. *Macular Edema*. 2010;47:73−110. Karger Publishers https://www.karger.com/Article/Abstract/320075. Accessed May 31, 2022.
4. Varma R, Bressler NM, Doan QV, et al. Prevalence of and risk factors for diabetic macular edema in the United States. *JAMA Ophthalmol*. 2014;132:1334−1340.
5. Klein R, Klein BEK, Moss SE, Cruickshanks KJ. The Wisconsin epidemiologic study of diabetic retinopathy XV. *Ophthalmology*. 1995;102:7−16.
6. White NH, Sun W, Cleary PA, et al. Effect of prior intensive therapy in type 1 diabetes on 10-year progression of

retinopathy in the DCCT/EDIC: comparison of adults and adolescents. *Diabetes*. 2010;59:1244−1253.
7. Anon. Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified Airlie house classification. *Ophthalmology*. 1991;98:786−806.
8. Diabetic Retinopathy Clinical Research Network, Elman MJ, Aiello LP. Randomized trial evaluating ranibizumab plus prompt or deferred laser or triamcinolone plus prompt laser for diabetic macular edema. *Ophthalmology*. 2010;117:1064−1077.e35.
9. Kim EJ, Lin WV, Rodriguez SM, et al. Treatment of diabetic macular edema. *Curr Diab Rep*. 2019;19:68.
10. Bhagat N, Grigorian RA, Tutela A, Zarbin MA. Diabetic macular edema: pathogenesis and treatment. *Surv Ophthalmol*. 2009;54:1−32.
11. International Diabetes Federation. *Clinical Practice Recommendations for Managing Diabetic Macular Edema*. Brussels, Belgium: International Diabetes Federation; 2019.
12. Panozzo G, Parolini B, Gusson E, et al. Diabetic macular edema: an OCT-based classification. *Semin Ophthalmol*. 2004;19:13−20.
13. Lewis K. Improving patient compliance with diabetic retinopathy screening and treatment. *Community Eye Health*. 2015;28:68−69.
14. Saadine JB, Fong DS, Yao J. Factors associated with follow-up eye examinations among persons with diabetes. *Retina*. 2008;28:195−200.
15. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. *J Glob Health*. 2018;8:020303.
16. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *NPJ digital medicine*. 2018;1:1−10.
17. Sierra-Sosa D, Garcia-Zapirain B, Castillo C, et al. Scalable healthcare assessment for diabetic patients using deep learning on multiple GPUs. *IEEE Trans Ind Inf*. 2019;15:5682−5689.
18. Alyoubi WL, Shalash WM, Abulkhair MF. Diabetic retinopathy detection through deep learning techniques: a review. *Inform Med Unlocked*. 2020;20:100377.
19. Grzybowski A, Brona P, Lim G, et al. Artificial intelligence for diabetic retinopathy screening: a review. *Eye*. 2020;34:451−460.
20. U.S. Food and Drug Administration. Center for drug evaluation and research. AEYE health K221183 approval letter. https://www.accessdata.fda.gov/cdrh_docs/pdf22/K221183.pdf. Accessed November 10, 2022.
21. Alsaih K, Lemaitre G, Rastgoo M, et al. Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. *Biomed Eng Online*. 2017;16:68.
22. Kaymak S, Serener A. *Automated age-related macular degeneration and diabetic macular edema detection on OCT images using deep learning*. New York, NY: IEEE; 2018.
23. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172:1122−1131.e9.
24. Manikandan S, Raman R, Rajalakshmi R, et al. Deep learning−based detection of diabetic macular edema using optical coherence tomography and fundus images: a meta-analysis. *Indian J Ophthalmol*. 2023;71:1783.
25. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402−2410.
26. Mo J, Zhang L, Feng Y. Exudate-based diabetic macular edema recognition in retinal images using cascaded deep residual networks. *Neurocomputing*. 2018;290:161−171.
27. Sahlsten J, Jaskari J, Kivinen J, et al. Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Sci Rep*. 2019;9:10750.
28. Li F, Wang Y, Xu T, et al. Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. *Eye*. 2021;36:1433−1441.
29. Yao Z, Yuan Y, Shi Z, et al. FunSwin: a deep learning method to analysis diabetic retinopathy grade and macular edema risk based on fundus images. *Front Physiol*. 2022;13:961386.
30. Varadarajan AV, Bavishi P, Ruamviboonsuk P, et al. Predicting optical coherence tomography-derived diabetic macular edema grades from fundus photographs using deep learning. *Nat Commun*. 2020;11:130.
31. Paul W, Burlina P, Mocharla R, et al. Accuracy of artificial intelligence in estimating best-corrected visual acuity from fundus photographs in eyes with diabetic macular edema. *JAMA Ophthalmol*. 2023;141:677−685.
32. Cuadros J, Bresnick G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. *J Diabetes Sci Technol*. 2009;3:509−516.
33. Bresnick GH, Mukamel DB, Dickinson JC, Cole DR. A screening approach to the surveillance of patients with diabetes for the presence of vision-threatening retinopathy. *Ophthalmology*. 2000;107:19−24.
34. Litvin TV, Ozawa GY, Bresnick GH, et al. Utility of hard exudates for the screening of macular edema. *Optom Vis Sci*. 2014;91:370−375.
35. Decencière E, Zhang X, Cazuguel G, et al. Feedback on a publicly distributed image database: the Messidor database. *Image Anal Stereol*. 2014;33:231.
36. Bressler I, Aviv R, Margalit D, et al. Autonomous screening for laser photocoagulation in fundus images using deep learning. *Br J Ophthalmol*; 2023. https://bjo.bmj.com/content/early/2023/05/22/bjo-2023-323376. Accessed June 6, 2023.
37. Bressler I, Aviv R, Margalit D, et al. Autonomous screening for laser photocoagulation in fundus images using deep learning. *Br J Ophthalmol*. 2024;108:742−746.
38. Rom Y, Aviv R, Ianchulev T, Dvey-Aharon Z. Predicting the future development of diabetic retinopathy using a deep learning algorithm for the analysis of non-invasive retinal imaging. *BMJ Open Ophthalmol*. 2022;7:e001140.
39. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York, NY: Chapman and Hall/CRC; 1994.
40. Porwal P, Pachade S, Kamble R, et al. Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research. *Data*. 2018;3:25.
41. Wong TY, Klein R, Islam FA, et al. Diabetic retinopathy in a multi-ethnic cohort in the United States. *Am J Ophthalmol*. 2006;141:446−455.
42. Selvaraju RR, Cogswell M, Das A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE international conference on computer vision;* 2017:618−626. http://openaccess.thecvf.com/content_iccv_2017/html/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.html. Accessed January 2, 2024.
43. Mansour SE, Browning DJ, Wong K, et al. The evolving treatment of diabetic retinopathy. *Clin Ophthalmol*. 2020;14:653−678.
44. Lee R, Wong TY, Sabanayagam C. Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss. *Eye Vis*. 2015;2:1−25.
45. Wang T-Y, Chen Y-H, Chen J-T, et al. Diabetic macular edema detection using end-to-end deep fusion model and

anatomical landmark visualization on an edge computing device. *Front Med*. 2022;9. https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2022.851644/full. Accessed July 22, 2024.

46. Food and Drug Administration (FDA). FDA device regulation: 510(k), K221183. 2022. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K221183. Accessed March 14, 2023.
47. Food and Drug Administration (FDA). FDA device regulation: 510(k), K240058. 2024. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K240058. Accessed June 26, 2024.
48. Date RC, Shen KL, Shah BM, et al. Accuracy of detection and grading of diabetic retinopathy and diabetic macular edema using teleretinal screening. *Ophthalmol Retina*. 2019;3:343–349.