OPEN ACCESS

Autonomous screening for laser photocoagulation in fundus images using deep learning

Idan Bressler ,[1] Rachelle Aviv ,[1] Danny Margalit,[1] Yovel Rom,[1] Tsontcho Ianchulev ,[1,2] Zack Dvey-Aharon[1]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bjo-2023-323376).

[1]AEYE Health, New York, New York, USA
[2]Ophthalmology, New York Eye and Ear Infirmary of Mount Sinai, New York, New York, USA

Correspondence to
Rachelle Aviv, AEYE Health, Tel Aviv 6473925, Israel; rachelle@aeyehealth.com

Received 12 February 2023
Accepted 15 April 2023

To cite: Bressler I, Aviv R, Margalit D, *et al*. *Br J Ophthalmol* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/bjo-2023-323376

ABSTRACT

Background Diabetic retinopathy (DR) is a leading cause of blindness in adults worldwide. Artificial intelligence (AI) with autonomous deep learning algorithms has been increasingly used in retinal image analysis, particularly for the screening of referrable DR. An established treatment for proliferative DR is panretinal or focal laser photocoagulation. Training autonomous models to discern laser patterns can be important in disease management and follow-up.

Methods A deep learning model was trained for laser treatment detection using the EyePACs dataset. Data was randomly assigned, by participant, into development (n=18 945) and validation (n=2105) sets. Analysis was conducted at the single image, eye, and patient levels. The model was then used to filter input for three independent AI models for retinal indications; changes in model efficacy were measured using area under the receiver operating characteristic curve (AUC) and mean absolute error (MAE).

Results On the task of laser photocoagulation detection: AUCs of 0.981, 0.95, and 0.979 were achieved at the patient, image, and eye levels, respectively. When analysing independent models, efficacy was shown to improve across the board after filtering. Diabetic macular oedema detection on images with artefacts was AUC 0.932 vs AUC 0.955 on those without. Participant sex detection on images with artefacts was AUC 0.872 vs AUC 0.922 on those without. Participant age detection on images with artefacts was MAE 5.33 vs MAE 3.81 on those without.

Conclusion The proposed model for laser treatment detection achieved high performance on all analysis metrics and has been demonstrated to positively affect the efficacy of different AI models, suggesting that laser detection can generally improve AI-powered applications for fundus images.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Laser photocoagulation is an established treatment for retinal conditions such as Diabetic Retinopathy. The performance of artificial intelligence models for the detection of various retinal indications may be affected by the existence of laser artifacts.

WHAT THIS STUDY ADDS

⇒ This study proposes a new state-of-the-art artificial intelligence model for the detection of laser photocoagulation artifacts, as well as demonstrating a positive effect on the performance of other artificial intelligence models.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study may improve or aid the development of future artificial intelligence models for the detection and diagnosis of various retinal conditions.

INTRODUCTION

Laser photocoagulation is a common and established procedure, in which laser pulses are used to coagulate retinal tissue, used to treat multiple retinal diseases.[1–3] Ablative photocoagulation is mostly used to prevent leakage and ischaemic neovascularisation in vascular retinal conditions such as diabetic retinopathy (DR),[4 5] diabetic macular oedema (DME),[6–8] retinal vein occlusion,[9 10] and neovascular age-related macular degeneration (AMD).[11]

Laser photocoagulation is generally divided into panretinal and focal; the former is delivered in the peripheral retina with deep ablative burns to stem the neovascular process,[12 13] while the latter is a lighter photocoagulative treatment delivered in the central macula to treat macular conditions.[14 15] There are well-established laser treatment protocols depending on disease severity and individual patient disease state.[11 16–18] While laser photocoagulation is an effective treatment, it causes retinal scarring and is destructive to the retinal tissue leaving long-term defects in the anatomy.[19–21]

Artificial intelligence (AI) using fundus imaging has been increasingly employed in various ophthalmological applications.[22 23] These applications include extraction of basic patient data, such as age and sex,[24] detection of retinal pathologies,[25 26] and pathology development prediction.[27 28] AI methods rely on image pattern recognition, especially in areas in which the pathology is present. As such, laser photocoagulation may disrupt general pattern recognition by adding new patterns or artefacts, such as burns and scars, which the model is less trained to deal with. This is specifically problematic given that laser treatment is often done on areas of interest, such as leaky blood vessels, which are often the very areas that are most crucial to recognise.

The effect laser photocoagulation has on AI systems suggests that a tool to identify images of eyes which have undergone photocoagulation may be beneficial for the autonomous retinal-based diagnosis and follow-up treatment of patients. While previous methods of laser photocoagulation

Table 1 Laser treatment prevalence in the EyePACs dataset

	No laser treatment	Focal laser	Panretinal laser	Both
Count	10525	1888	9484	847

Table 2 Patient demographics for patients who did and did not have laser treatment artefacts

	Count	Age (SD)	Gender (% female)	Ethnicity (fraction)
With laser	10525	59.5 (10.0)	51	White=0.59 (Hispanic=0.95 non-Hispanic=0.05) Indian subcontinent origin=0.11 African Descent=0.07 Asian=0.02 Not specified=0.13 Other=0.08
No laser	10525	54.7 (10.8)	52	White=0.51 (Hispanic=0.90 non-Hispanic=0.10) African Descent=0.13 Indian subcontinent origin=0.12 Asian=0.03 Not specified=0.14 Other=0.07

detection exist,[29–33] this work, to the best of our knowledge, presents the first laser treated image detection method based on a large, diverse, and widely accepted database—in this case, EyePACS (https://www.eyepacs.org); the database contains images from a variety of manufacturers and patient populations, of varying image qualities.

METHODS

Data

The data consisted of a subsample of the EyePACs dataset, which contains 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.

The dataset contained up to six images per patient visit: one macula centred image, one disc centred image, and one centred image (in which a central fixation image is fixated on the middle of a line connecting the foveola and the optic disc), per eye. Each eye underwent expert reading, including but not limited to panretinal laser treatment presence, focal laser treatment presence, and image quality. All images of the subsample deemed readable by expert annotations were used.

The resulting dataset consisted of 21050 images from 9212 patients, of which 9484 images (45%) had artefacts of panretinal laser treatment, 1888 (9%) had artefacts of focal laser treatment and 847 (4%) had both. This work combined focal and panretinal laser treatments into 1 category of laser treatment, resulting in an overall 10525 (50%) images with laser treatment artefacts (table 1). Of these, roughly 77% of patients required dilation, where 54% of all patients received 1 gtt. tropicamide 1%, 17% received 1 gtt. tropicamide 0.5% and 5% received other dilation agents.

The average age of patients with laser treatment artefacts was 59.5 (10.0 SD) and 55% were women, compared with the patients who had not undergone laser treatment, for which the average age was 55.6 (11.3 SD) and 61% of which were women (table 2). The prevalence of laser photocoagulation across ethnic groups may be found in online supplemental table A. The distribution of laser treatment images across DR levels is given in online supplemental table B; all laser treatment images were from patients with more than mild DR, and the majority were from patients with grade 4 DR.

Quality assessment

An image quality assessment tool was developed using classic computer vision methods; the tool detects visibility of fundus-specific characteristics and assigns each image a score. The given quality score for an image is an aggregation of the visibility from multiple areas within the fundus image. The tool was validated based on visual assessment of images score and the readability of the images. Figure 1 demonstrates a few examples of images and their respective scores, showing the correlation between score and visual image quality. This was done in order to remove low-quality images from the dataset, as the quality scores assigned by EyePACs are assigned to patients and not to individual images.

Preprocessing

Image preprocessing was performed in two steps for both datasets. First, image backgrounds were cut along the convex hull, which contains the circular border between the image and the background. Figure 2 shows an example of this process. Second, images were resized to 512×512 pixels. Lastly, using the afore-mentioned quality assessment tool, bad-quality images were filtered out before training. The model was checked with multiple training configurations set by multiple thresholds and the image quality threshold was set at the point at which model performances were not improved by filtering additional images, resulting in 1373 images filtered, approximately 6.5% of the data.

Model training

The data was then divided into training, validation, and test datasets at a ratio of 80%, 10% and 10%, respectively. A binary

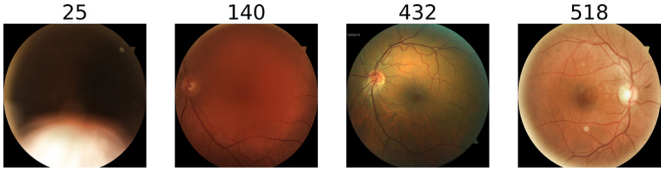

Figure 1 Example images and their accompanied image quality scores, ordered from the worst quality (left) to the best quality (right).

Figure 2 Example of image cropping, blue lines represent the cropping boundaries.

Table 3 Laser treatment detection results on the EyePACs dataset for the three analysis levels performed, given in accuracy, sensitivity, specificity and AUC with a 95% CI

	Accuracy (CI)	Sensitivity (CI)	Specificity (CI)	AUC (CI)
Image level	0.882 (0.870 to 0.892)	0.883 (0.868 to 0.897)	0.880 (0.864 to 0.894)	0.950 (0.943 to 0.956)
Eye level	0.929 (0.916 to 0.940)	0.925 (0.900 to 0.945)	0.931 (0.916 to 0.944)	0.979 (0.972 to 0.984)
Patient level	0.927 (0.910 to 0.940)	0.929 (0.881 to 0.947)	0.926 (0.911 to 0.944)	0.981 (0.971 to 0.987)

CI noted in parentheses.
AUC, area under the receiver operating characteristic curve.

classification neural network was trained. The model architecture was automatically fitted to best balance the model performance versus model complexity tradeoff. Hyperparameter tuning was done on the validation set.

Statistical analysis

The metrics used for model assessment were accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). For each metric, the bias corrected and accelerated bootstrap method[34] was used to produce a 95% CI.

Analysis levels

Laser detection was done on three different levels. The first, detection on the individual image level, was the basic task for which the model was trained. The second, detection on the eye level, used all images from a given eye and the image for which the model had the highest probability score was selected for analysis. For the third, detection on the patient level, the results from both eyes were compared and the eye with the higher probability score was selected to produce a patient-level result. The eye and patient level analysis, respectively, operate on the logic that one field or eye with photocoagulation artefacts is sufficient for the eye or patient to be classified as positive.

Effect on imaging tasks

The effect that laser treatment has on imaging tasks was measured by applying the laser detection model as a postprocessing step for a model for the detection of DME, which was developed based on the EyePACs dataset,[35] and a model for age detection, also developed based on the EyePACs dataset.

The performance on these tasks was measured in AUC on a separate validation set containing images both with and without laser treatment artefacts. The 95% CI was calculated using the accelerated bootstrap method for each population and compared for significance.

A regression model was additionally trained for age detection, and the mean absolute error (MAE) between the patient's age and predicted age was calculated on a separate validation set. The validation set was separated into patients with and without laser treatment artefacts, such that the mean age between these populations was the same. Significance in MAE between the two populations was calculated using a student's t-test. Detailed patient statistics of these experiments, as well as details on model development, are given in online supplemental tables C and following explanations.

RESULTS

The results for the different analysis methods of laser artefact detection were as follows (table 3): on the image level, sensitivity of 0.883 (95% CI 0.868 to 0.897), specificity of 0.880 (95% CI 0.864 to 0.894), and AUC of 0.950 (95% CI 0.943 to 0.956) were achieved. On the eye level, sensitivity of 0.925 (95% CI 0.900 to 0.945), specificity of 0.931 (95% CI 0.916 to 0.944), and AUC of 0.979 (95% CI 0.972 to 0.984) were achieved. On the patient level, sensitivity of 0.929 (95% CI 0.881 to 0.947), specificity of 0.926 (95% CI 0.911 to 0.944), and AUC of 0.981 (95% CI 0.971 to 0.987) were achieved.

The results of laser artefact detection for each DR level are displayed in table 4: the model achieved 0.910 AUC (95% CI 0.866 to 0.941) for DR level 2, 0.887 AUC (95% CI 0.758 to 0.954) for DR level 3, 0.929 AUC (95% CI 0.918 to 0.938) for DR level 4, and 0.772 AUC (95% CI 0.904 to 0.968) for ungradable DR level. DR levels 0 and 1 did not have any laser treated examples, thus most metrics are not defined for these groups. The results of laser artefact detection stratified by ethnicity are available in online supplemental table D.

Online supplemental table E shows the difference in results in laser artefact detection between patients with and without DME. The model achieved 0.955 AUC (0.948–0.962) for non DME patients versus 0.908 AUC (0.884–0.927) for DME patients, demonstrating that these conditions do affect results, but the model achieves high performance irrespective of them.

Online supplemental table F displays the results of laser artefact detection for images which passed (high quality) and did not pass (low quality) the quality filter, showing a significant difference between the populations. The results for low-quality images, which were filtered out, were 0.787 sensitivity (95% CI 0.710 to 0.849), 0.793 specificity (95% CI 0.709 to 0.860), and 0.857 AUC (95% CI 0.803 to 0.898); compared with 0.854 sensitivity (95% CI 0.838 to 0.869), 0.904 specificity (95% CI 0.890 to 0.917), and 0.948 AUC (95% CI 0.941 to 0.955) for high-quality images which passed the filter.

The effect of laser detection and subsequent filtration on the afore-mentioned three tasks of DME detection, age prediction,

Table 4 Results on the EyePACs dataset across DR grades, given in accuracy, sensitivity, specificity, and AUC with a 95% CI

DR grade	2	3	4	Ungradable
Sensitivity (CI)	0.680 (0.591 to 0.756)	0.667 (0.472 to 0.806)	0.869 (0.853 to 0.884)	0.846 (0.652 to 0.957)
Specificity (CI)	0.958 (0.906 to 0.984)	0.960 (0.791 to 1)	0.845 (0.821 to 0.866)	0.688 (0.400 to 0.882)
AUC (CI)	0.910 (0.866 to 0.941)	0.887 (0.758 to 0.954)	0.929 (0.918 to 0.938)	0.772 (0.904 to 0.968)

CI noted in parentheses.
DR, diabetic retinopathy.

Table 5 Results for the three experiments conducted for the effect of laser treatment, showing the results in terms of AUC for sex and DME detection and mean average error for age detection.

	Metric for laser images	Metric for no laser images	% Laser treated
DME	0.932 (95% CI 0.905 to 0.951) AUC	0.955 (95% CI 0.948 to 0.961) AUC	12.5%
Sex	0.872 (95% CI 0.830 to 0.903) AUC	0.922 (95% CI 0.916 to 927) AUC	3%
Age	3.81 MAE	5.33 MAE	5%

The percentage of filtered images is shown.
AUC, area under the receiver operating characteristic curve; DME, diabetic macular oedema; MAE, mean absolute error.

and sex detection were as follows: DME detection results for images with no laser artefacts were 0.955 AUC (95% CI to 0.948 to 0.961), compared with images with laser artefacts, on which the model achieved 0.932 AUC (95% CI 0.905 to 0.951). Age prediction results for images with no laser artefacts, after age adjustment, were 3.81 MAE, compared with images with laser artefacts, on which the model achieved 5.33 MAE. T-test analysis shows a significance of $p<1e-4$. Sex detection results for images with no laser artefacts were 0.922 AUC (95% CI 0.916 to 927), compared with images with laser artefacts for which the model achieved 0.872 AUC (95% CI 0.830 to 0.903).

The aggregation of these results is shown in table 5.

DISCUSSION

This work proposed a method for the automatic detection of laser treatment artefacts in fundus images, which may also serve as a component in the future development of AI systems for different diagnoses based on retinal imaging. Such tasks may need to consider images of laser-treated eyes differently from non-treated eyes according to their design needs; some may choose to discard these images, while others may analyse them in a manner differently to images of untreated eyes. Accordingly, and in accordance with the degree to which laser treatment affects the task in question, the proposed system may be used at different operating points with different sensitivity–specificity balances. Discarding laser-treated images is a viable option for most automated retinal screening applications, as these patients should already have an awareness of the need for regular screening.

Previous studies on the autonomous detection of laser burns from fundus images have been on a smaller scale (roughly 2 orders of magnitude).[29–33] The importance of scale is in the better representation of real-life conditions; specifically, this study allows better representation of various image qualities, camera manufacturers, and populations. Additionally, a wider range of clinical conditions, such as DR and DME, are represented in this study both with and without laser treatment, and the proposed system shows high performance across these conditions.

The effect laser treatment has on imaging tasks, and the model's ability to detect relevant images was validated by checking the model's effect on different AI tasks involving retinal images. A significant difference was found for all three tasks, showing the relevance of the proposed method for future AI tasks.

A limitation of this work is the lack of differentiation between focal and panretinal laser treatments that were grouped as one in this work. Future works may differentiate between the two, given increased data. Furthermore, even though the base characteristics of laser photocoagulation remain similar across conditions, the addition of AMD-specific databases to the training set may improve results.

In addition, and in the same vein of the presented work, machine learning methods to detect patients with DME who will require future laser treatment may be developed. This would require training a model, similar to the one presented, on a dataset generated from a longitudinal study tracking the progression of patients with diabetes.

Contributors IB analysed the data, designed the study and conducted research. ZD-A and DM conceived the study and supervised research. YR provided assistance in assessing external models. TI provided medical and strategic guidance and oversight. IB and RA drafted the manuscript with input from all authors. ZD-A is guarantor.

Funding Employees and board members of AEYE Health designed and carried out the study; managed, analysed and interpreted the data; prepared, reviewed and approved the article; and were involved in the decision to submit the article. There were no grants or awards involved in the funding of this article.

Competing interests RA, IB and YR are employees of AEYE Health. TI is a board member of AEYE Health. DM is COO of AEYE Health. ZD-A is CEO of AEYE Health.

Patient consent for publication Not applicable.

Ethics approval Institutional Review Board exemption was obtained from the Sterling Independent Review Board on the basis of a category 4 exemption (DHHS), pursuant to the terms of the U.S. Department of Health and Human Service's Policy for Protection of Human Research Subjects at 45 C.F.R §46.104(d)

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Deidentified data used in this study are not publicly available at present. Parties interested in data access should contact Jorge Cuadros (jcuadros@eyepacs.com) for queries related to EyePACS.

ORCID iDs
Idan Bressler http://orcid.org/0000-0002-5058-2307
Rachelle Aviv http://orcid.org/0000-0002-0350-9238
Tsontcho Ianchulev http://orcid.org/0000-0002-9893-5909

REFERENCES

1 Krauss JM, Puliafito CA. Lasers in Ophthalmology. *Lasers Surg Med* 1995;17:102–59.
2 Kozak I, Luttrull JK. Modern retinal laser therapy. *Saudi J Ophthalmol* 2015;29:137–46.
3 Lock JH, Fong KCS, Lock JH. Retinal laser Photocoagulation. *Med J Malaysia* 2010;65:88–94.
4 Neubauer AS, Ulbig MW. Laser treatment in diabetic retinopathy. *Ophthalmologica* 2007;221:95–102.
5 Evans JR, Michelessi M, Virgili G. Laser Photocoagulation for proliferative diabetic retinopathy. *Cochrane Database Syst Rev* 2014;2014:CD011234.
6 Romero-Aroca P, Reyes-Torres J, Baget-Bernaldiz M, *et al*. Laser treatment for diabetic macular edema in the 21st century. *Curr Diabetes Rev* 2014;10:100–12.
7 Mitchell P, Wong TY, Diabetic Macular Edema Treatment Guideline Working Group. Management paradigms for diabetic macular edema. *Am J Ophthalmol* 2014;157:505–13.
8 Jorge EC, Jorge EN, Botelho M, *et al*. Monotherapy laser Photocoagulation for diabetic macular oedema. *Cochrane Database Syst Rev* 2018;10:CD010859.
9 Wong TY, Scott IU. Retinal-vein occlusion. *N Engl J Med* 2010;363:2135–44.

10 Rehak M, Wiedemann P. Retinal vein thrombosis: Pathogenesis and management. *J Thromb Haemost* 2010;8:1886–94.
11 Querques G, Cicinelli MV, Rabiolo A, *et al*. Laser Photocoagulation as treatment of non-Exudative age-related macular degeneration: State-of-the-art and future perspectives. *Graefes Arch Clin Exp Ophthalmol* 2018;256:1–9.
12 Bressler NM, Beck RW, Ferris FL. Panretinal photocoagulation for proliferative diabetic retinopathy. *N Engl J Med* Overseas Ed 2011;365:1520–6.
13 Royle P, Mistry H, Auguste P, *et al*. Pan-Retinal photocoagulation and other forms of laser treatment and drug therapies for non-proliferative diabetic retinopathy: systematic review and economic evaluation. *Health Technol Assess* 2015;19:1–248.
14 Beck RW, Edwards AR, Aiello LP, *et al*. Three-year follow-up of a randomized trial comparing focal/grid Photocoagulation and intravitreal triamcinolone for diabetic macular edema. *Arch Ophthalmol* 2009;127:245–51.
15 ETDRSR Group. Treatment techniques and clinical guidelines for Photocoagulation of diabetic macular edema: Early treatment diabetic retinopathy study report number 2. *Ophthalmology* 1987;94:761–74.
16 ETDRSR Group. Early Photocoagulation for diabetic retinopathy ETDRS report. *Ophthalmology* 1991;98:766–85.
17 Focal photocoagulation treatment of diabetic macular edema. Relationship of treatment effect to fluorescein angiographic and other retinal characteristics at baseline: ETDRS report no. 19. early treatment diabetic retinopathy study research group. *Arch Ophthalmol Chic Ill 1960* 1995;113:1144–55.
18 McIntosh RL, Mohamed Q, Saw SM, *et al*. Interventions for central retinal vein occlusion: An evidence-based systematic review. *Ophthalmology* 2007;114:835–54.
19 Maeshima K, Utsugi-Sutoh N, Otani T, *et al*. Progressive enlargement of scattered Photocoagulation scars in diabetic retinopathy. *Retina* 2004;24:507–11.
20 Schatz H, Madeira D, McDonald HR, *et al*. Progressive enlargement of laser scars following grid laser Photocoagulation for diffuse diabetic macular edema. *Arch Ophthalmol* 1991;109:1549–51.
21 Greenstein VC, Chen H, Hood DC, *et al*. Retinal function in diabetic macular edema after focal laser Photocoagulation. *Invest Ophthalmol Vis Sci* 2000;41:3655–64.
22 Li T, Bo W, Hu C, *et al*. Applications of deep learning in fundus images: a review. Med Image Anal 2021;69:101971.
23 Jeong Y, Hong Y-J, Han J-H. Review of machine learning applications using retinal fundus images. *Diagnostics (Basel)* 2022;12:134.
24 Poplin R, Varadarajan AV, Blumer K, *et al*. Prediction of cardiovascular risk factors from retinal fundus Photographs via deep learning. Nat Biomed Eng 2018;2:158–64.
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;316:2402–10.
26 et alChen X, Xu Y, Kee Wong DW. Glaucoma detection based on deep Convolutional neural network. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan. IEEE, 2015:715–8.
27 Bora A, Balasubramanian S, Babenko B, *et al*. Predicting the risk of developing diabetic retinopathy using deep learning. *Lancet Digit Health* 2021;3:e10–9.
28 Rom Y, Aviv R, Ianchulev T, *et al*. Predicting the future development of diabetic retinopathy using a deep learning algorithm for the analysis of non-invasive retinal imaging. *BMJ Open Ophth* 2022;7:e001140.
29 Dias JMP, Oliveira CM, da Silva Cruz LA. Detection of laser marks in retinal images.
30 Sousa JGA, Oliveira CM, Silva Cruz LA. Automatic detection of laser marks in retinal Digital fundus images. IEEE, 2016:1313–7
31 Tahir F, Akram MU, Abbass M. Laser marks detection from fundus images. 2014 14th International Conference on Hybrid Intelligent Systems (HIS); Kuwait. IEEE, 2014:147–51
32 Lam C, Yu C, Huang L, *et al*. Retinal lesion detection with deep learning using image patches. *Invest Ophthalmol Vis Sci* 2018;59:590–6.
33 Pan X, Jin K, Cao J, *et al*. Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning. *Graefes Arch Clin Exp Ophthalmol* 2020;258:779–85.
34 Efron B, Tibshirani RJ. An introduction to the Bootstrap. CRC Press, 1994.
35 Bressler I, Aviv R, Margalit D, *et al*. Autonomous screening for diabetic macular edema using deep learning processing of retinal images. *Ophthalmology* [Preprint].

Supplementary

Table A: Ethnic distribution of training dataset

	Population	With Photocoagulation	No Photocoagulation
Latin American	0.52	0.56	0.47
Not Specified	0.14	0.13	0.14
Indian subcontinent origin	0.12	0.11	0.12
African Descent	0.10	0.07	0.13
Caucasian	0.04	0.03	0.05
Asian	0.03	0.03	0.03
Others	0.05	0.07	0.06

Bressler I, *et al. Br J Ophthalmol* 2023;0:1–5. doi: 10.1136/bjo-2023-323376

Table B: Distribution of laser treatment prevalence across different diabetic retinopathy grades

DR grade	0	1	2	3	4	ungradable
Total count	558,451	52,928	106,377	16,909	19,470	25,718
With laser	0	0	365	175	9,607	136
No laser	558,451	52,928	105,742	16,734	9,863	25,582

Bressler I, *et al. Br J Ophthalmol* 2023;0:1–5. doi: 10.1136/bjo-2023-323376

Table C: Cohort statistics for the three modules tested for laser treatment filtering effect.

Task	Patients	Images	Mean Age (S.D)	Gender	Ethnicity (fraction)
DME detection	1,607	3,576	54.80 (10.35)	50% female	white = 0.55 (Hispanic = 0.93, non-Hispanic = 0.07) ethnicity not specified = 0.13 Indian subcontinent origin = 0.12 African Descent = 0.11 Asian = 0.02 Other = 0.07
Age	2612	4754	61.78 (10.88)	57% female	white = 0.50 (Hispanic = 0.88, non-Hispanic = 0.12) ethnicity not specified = 0.17 African Descent = 0.13 Indian subcontinent origin = 0.08 Asian = 0.03 Other = 0.09
Sex	1318	10506	54.94 (11.06)	59% female	white = 0.60 (Hispanic = 0.83, non-Hispanic = 0.17) ethnicity not specified = 0.14 African Descent = 0.09 Indian subcontinent origin = 0.06 Asian = 0.03 Other = 0.08

Bressler I, *et al. Br J Ophthalmol* 2023;0:1–5. doi: 10.1136/bjo-2023-323376

Table C1: Calibration data cohort statistics for the three modules tested for laser treatment filtering effect.

Task	Patients	Images	Mean Age (S.D)	Gender	Ethnicity (fraction)
DME detection	1,607	3,576	54.80 (10.35)	50% female	white = 0.55 (Hispanic = 0.93, non-Hispanic = 0.07) ethnicity not specified = 0.13 Indian subcontinent origin = 0.12 African Descent = 0.11 Asian = 0.02 Other = 0.07
Age	9,485	76,173	55.18 (11.67)	59% female	white = 0.61 (Hispanic = 0.87, non-Hispanic = 0.13) ethnicity not specified = 0.18 African Descent = 0.10 Indian subcontinent origin = 0.04 Asian = 0.04 Other = 0.07
Sex	1,930	15,786	54.47 (11.37)	60% female	white = 0.66 (Hispanic = 0.85, non-Hispanic = 0.15) ethnicity not specified = 0.14 African Descent = 0.08 Indian subcontinent origin = 0.05 Asian = 0.05 Other = 0.02

Table C2: Test data cohort statistics for the three modules tested for laser treatment filtering effect.

Task	Patients	Images	Mean Age (S.D)	Gender	Ethnicity (fraction)
DME detection	14,078	28,274	55.00 (10.18)	51% female	white = 0.55 (Hispanic = 0.93, non-Hispanic = 0.07) ethnicity not specified = 0.13 Indian subcontinent origin = 0.11 African Descent = 0.11 Asian = 0.03 Other = 0.07
Age	85,374	686,255	55.17 (11.62)	58% female	white = 0.61 (Hispanic = 0.87, non-Hispanic = 0.13) ethnicity not specified = 0.18 African Descent = 0.10 Indian subcontinent origin = 0.05 Asian = 0.04 Other = 0.06
Sex	91,961	742,915	55.23 (11.64)	59% female	white = 0.63 (Hispanic = 0.87, non-Hispanic = 0.13) ethnicity not specified = 0.15 African Descent = 0.1 Indian subcontinent origin = 0.05 Asian = 0.04 Other = 0.03

All models were trained using data compiled and provided by EyePacs which consisted of 45° angle fundus photography images and expert readings of said images. Images were resized to 512X512 pixels for all models.

DME model

The data was composed of 32,049 images from 15,892 patients. Two images were taken for each eye from two different fields, one centered on the macula and another centered between the macula and the disc. The average age was 55.02 (10.21 SD), 51% of the patients were female.

The data was divided into training, validation, and test datasets consisting of 80%, 10%, and 10% of the data respectively.

Sex model

The training & validation data was comprised of 742,915 images from 91,961 patients. Test data was comprised of 15,786 images from 1,930 patients.

Age model

The training & validation data was comprised of 686,225 images from 85,374 patients. Test data was comprised of 76,173 images from 9,485 patients.

Table D: laser artifact detection stratified by ethnicity.

	Accuracy	Sensitivity	Specificity	AUC	Support
Latin American	0.914	0.901	0.931	0.973	10847
Ethnicity not specified	0.909	0.877	0.94	0.963	2855
African Descent	0.864	0.778	0.912	0.936	2196
Indian subcontinent origin	0.832	0.776	0.882	0.905	2473
Caucasian	0.913	0.864	0.943	0.97	905
Asian	0.926	0.928	0.925	0.979	664
Multi-racial	0.888	0.889	0.887	0.962	160
Other	0.928	0.938	0.915	0.983	207

Bressler I, *et al. Br J Ophthalmol* 2023;0:1–5. doi: 10.1136/bjo-2023-323376

Table E: Results on the EyePACs dataset for patients with and without DME, given in accuracy, sensitivity, specificity, and AUC with a 95% confidence interval. CI noted in parentheses.

DME	No	Yes
Sensitivity (C.I)	0.873 (0.856, 0.889)	0.770 (0.727, 0.811)
Specificity (C.I)	0.906 (0.890, 0.919)	0.897 (0.962, 0.925)
AUC (C.I)	0.955 (0.948, 0.962)	0.908 (0.884, 0.927)

Table F: Results for images which were filtered out and not filtered out by the image quality tool, given in accuracy, sensitivity, specificity, and AUC with a 95% confidence interval.

	Sensitivity (C.I)	Specificity (C.I)	AUC (C.I)
Filtered out	0.787 (0.710, 0.849)	0.793 (0.709, 0.860)	0.857 (0.803, 0.898)
Remained	0.854 (0.838, 0.869)	0.904 (0.890, 0.917)	0.948 (0.941, 0.955)