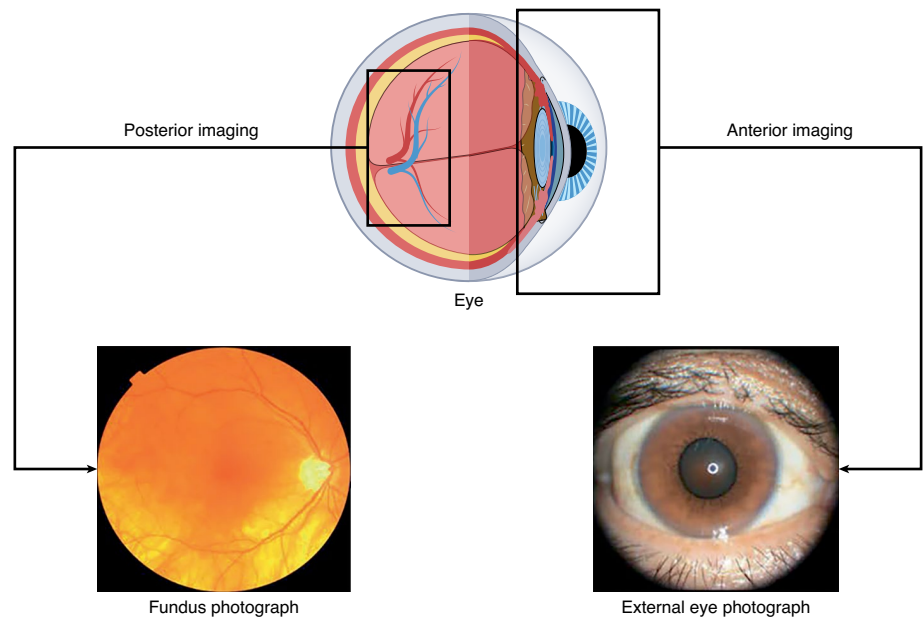MACHINE LEARNING

Eyeing severe diabetes upfront

Deep-learning models trained with images of the external part of the eyes, rather than fundus images of the retina, can also be used to detect severe diabetic conditions, such as diabetic retinopathy.

Zhen Ling Teo and Daniel Shu Wei Ting

Deep-learning models have been developed to detect tuberculosis from chest X-rays[1], malignant melanoma from skin photographs[2], breast cancer metastasis from histological sections[3] and, most recently, COVID-19 from computed-tomography images[4]. However, widespread clinical implementation of deep learning is hampered by diagnostic accuracy and by non-technical barriers, from patient engagement to administrative difficulties to ethical, privacy and accountability concerns[5] (some of these constraints to the clinical use of deep learning have, however, been alleviated during the COVID-19 pandemic[6]).

Beyond diagnostic accuracy, there are also a number of less obvious technical barriers to the clinical implementation of deep learning. This is exemplified by the development and testing of deep-learning systems for the screening of diabetic retinopathy (DR), a leading cause of preventable blindness. A universally accepted strategy to prevent blindness involves the identification of patients with poor glycaemic control for preventive care, the screening of patients for DR coupled with prompt referral, and the treatment of patients with advanced stages of DR[7]. Many deep-learning systems have shown excellent diagnostic performance for the detection of DR from images of the retina[8–10]. Notably, one such system aided the screening of patients with DR and reduced unnecessary tertiary referrals[11]. However, the system was implemented within a suitable screening programme using expensive specialized equipment (non-mydriatic fundus cameras), trained human graders and suitable software infrastructure. These types of constraint can hinder the development and applicability of deep-learning systems for patient screening, especially in low-to-middle-income countries[12].

Writing in *Nature Biomedical Engineering*, Yun Liu, Naama Hammel and co-authors now show that deep-learning models trained with photographs of the anterior part of the eyes (Fig. 1) can also be used to detect severe diabetic

Fig. 1 | Features of the front of the eye carry signs of diabetic complications. Complications of diabetes can be clinically diagnosed, also aided by machine learning, through fundus photography of the posterior segment of the eye. Photographs of the anterior part of the eye from standard consumer-grade cameras can also be used to detect diabetic conditions affecting the eyelids, conjunctiva, cornea and lens via deep-learning models trained on these images. Figure reproduced with permission from ref. [13], Springer Nature Ltd.

conditions — in particular, moderate non-proliferative DR (NPDR), vision-threatening DR (VTDR), diabetic macular oedema (DME) and poor glycaemic control[13]. The models were trained on eye photographs from 145,832 patients with diabetes (collected in the EyePACS database, which involves 301 DR screening sites in California), and tested on four independent datasets involving 48,644 patients from 198 additional screening sites across the United States (sites that participated in the EyePACS datasets located in states other than in California, and sites from the Diabetic Teleretinal Screening programme and Technology-Based Eye Care Services of the Atlanta Veterans Affairs healthcare system). For all patients, the datasets included photographs of the external part of their

eyes as well as retinal fundus photographs (taken according to the standard imaging protocol for DR screening), which had been graded for diabetic retinal disease by EyePACS-certified graders (for the EyePACS datasets) and by ophthalmologists (for the Atlanta Veterans Affairs datasets) on the basis of a suitable grading protocol (from the modified Early Treatment Diabetic Retinopathy Study[14]), and hence served as ground truth. Two of the four validation datasets contained photographs of the external part of the eyes taken after pupillary dilation; hence, to enhance the models' generalizability, the authors trained a separate deep-learning model for segmentation of the pupil and iris in the images. For this, 12 ophthalmologists graded 5,000 images (4,000 in the development set,

and 500 in each validation set) by drawing ellipses around each pupil and iris.

Liu and co-authors show that the trained deep-learning models predicted glycated haemoglobin (HbA1c) of ≥9% — a standard measure of average blood sugar levels in the past three months — with an area under the receiver operating characteristic curve (AUC) in the range 67.6–73.4%, as well as moderate NPDR, DME and VTDR with AUCs of 75–84%, 77.9–84.7% and 79.2–86.7%, respectively. The accuracy was better for photographs of dilated pupils. To extract explainability from the results of the trained models, the authors used heatmaps generated via ablation and saliency analyses. In the ablation analyses, the images were divided according to concentric circles (around the contours of pupils and irises) to mask either the circular region in the centre or the peripheral rim. In the saliency analyses, the images were first manually inspected to ensure their correct orientation. The analyses allowed the authors to identify key image features that contributed to the predictions of the models. Features in the nasal and temporal conjunctivas, where vessels are prominent, appeared to contribute the most to the predictions of HbA1c levels. The pupil region was more relevant for predicting moderate NPDR and DME (here, the models may leverage information from the light that reflects from the retina on the pupil rather than pupil size, as the authors postulate). Knowledge of these features can help improve the acquisition of eye photographs and model training, potentially improving model performance.

The study has several strengths. First, that photographs of the anterior region of the eye, captured by five types of camera, can be used to predict DR and glycaemic-control status without the need for posterior-segment examination or imaging is notable and could improve DR-screening services, especially in under-resourced countries. Second, the deep-learning models were developed using real-world image datasets from 301 DR-screening sites and tested on datasets from 198 additional sites, and the images had been assessed by graders trained on specific imaging protocols and classification systems. Third, the ablation and saliency analyses for explainability, and the sensitivity analyses for evaluating the impact of the minimally required resolution threshold for images with suboptimal quality, support the robustness of the deep-learning system and their use will be essential if the system is to be tested prospectively in real-world settings.

However, the study also highlights various challenges that would need to be addressed before eventual clinical implementation of the deep-learning system. First, sensitivity and specificity analyses using pre-set operating thresholds, tailored to the intended use environment, would need to be carried out. With its current accuracy (AUCs in the range 75–87%), the system is unlikely to replace the well-established DR-screening protocols currently performed by human graders or by human graders aided by deep-learning-based retinal imaging[9,10,15]. Second, the system may not be readily adopted by general practitioners or primary eye-care providers, who are used to performing posterior-segment photography with fundus cameras to diagnose DR. Third, although Liu and co-authors claim that the deep-learning system could be used as an early alert system to prompt patients when their HbA1c is >9%, these patients are likely to have already been identified by physicians as being at high risk for poorly controlled HbA1c and would thus be undergoing regular check-ups. Fourth, it is unclear whether the deep-learning system would maintain accuracy when using eye photographs taken with smartphone cameras under varying lighting conditions.

With an eye towards clinical implementation, further developments of the deep-learning system will need to involve the identification of the system's intended users (patients or primary care providers) and use settings (a hospital setting or at-home use via a smartphone app), determination of the point of deployment (smartphones or retinal cameras), pre-setting the operating thresholds for real-world deployment, testing the system's generalizability using other forms of photography (such as smartphone cameras), engagement of primary care providers or endocrinologists to design the clinical and patient-referral workflows, and a health-economics analysis. Notwithstanding the implementation challenges, Liu and colleagues' study makes it clear that eye photographs contain useful information for the detection of systemic diseases and, hence, that it is valuable to further explore the eye as a window to the body's health. ❐

Zhen Ling Teo[1] and
Daniel Shu Wei Ting [iD][1,2] ✉

[1]Singapore National Eye Centre, Singapore Eye Research Institute, Singapore, Singapore. [2]Duke–NUS Medical School, Singapore, Singapore.
✉e-mail: daniel.ting@duke-nus.edu.sg

References

1. Lakhani, P. & Sundaram, B. *Radiology* **284**, 574–582 (2017).
2. Esteva, A. et al. *Nature* **542**, 115–118 (2017).
3. Ehteshami Bejnordi, B. et al. *JAMA* **318**, 2199–2210 (2017).
4. Ning, W. *Nat. Biomed. Eng.* **4**, 1197–1207 (2020).
5. Davenport, T. & Kalakota, R. *Future Health. J.* **6**, 94–98 (2019).
6. Ting, D. S. W. et al. *Nat. Med.* **26**, 459–461 (2020).
7. Wong, T. Y. et al. *Ophthalmology* **125**, 1608–1622 (2018).
8. Abramoff, M. D. et al. *Invest. Ophthalmol. Vis. Sci.* **57**, 5200–5206 (2016).
9. Gulshan, V. et al. *JAMA* **316**, 2402–2410 (2016).
10. Ting, D. S. W. et al. *JAMA* **318**, 2211–2223 (2017).
11. Xie, Y. et al. *Lancet Digit. Health* **2**, e240–e249 (2020).
12. Bellemo, V. et al. *Lancet Digit. Health* **1**, e35–e44 (2019).
13. Babenko, B. et al. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-022-00867-5 (2022).
14. Early Treatment Diabetic Retinopathy Study Research Group. *Ophthalmology* **98**, 786–806 (1991).
15. Abramoff, M. D. et al. *npj Digital Med.* **1**, 39 (2018).