BMC Ophthalmology

RESEARCH

Open Access

Ensemble machine learning algorithm for anti-VEGF treatment efficacy prediction in diabetic macular edema

Yu Fang[1†], Jianwei Lin[1†], Peiwen Xie[1], Huishan Zhu[1], Tsz Kin Ng[1] and Guihua Zhang[1*]

Abstract

Background Diabetic macular edema (DME) is a leading cause of vision loss in diabetes, with variable responses to anti-vascular endothelial growth factor (anti-VEGF) therapy in DME patients. Current diagnosis relies on optical coherence tomography (OCT) imaging, but manual interpretation is limited. This study aims to integrate 3D-OCT features and clinical variables to develop machine learning (ML) models for predicting anti-VEGF treatment outcomes.

Methods and analysis Medical records and 3D-OCT images of DME patients were included in this study. The 3D-OCT images were categorized into good and poor visual response groups based on the best corrected visual acuity at one month after three consecutive anti-VEGF treatments. The images and clinical features were subjected to assessment by 11 automatic classification models for anti-VEGF treatment responses in DME patients. The top 3 performing models were selected to build an ensemble model, and evaluated in the test dataset.

Results This study included 142 patients with 3D-OCT images of 170 eyes. A total of 20 image and clinical features were selected for the model construction and test in DME patients responded to anti-VEGF therapy. Adaptive boosting (AdaBoost), GradientBoosting, and light gradient boosting machine (LightGBM) exhibited better performances than the remaining 8 models. The ensemble model constructed achieved a sensitivity of 0.941, specificity of 0.882, and accuracy of 0.912 in the test dataset, with an area under the receiver operating characteristic curve of 0.976.

Conclusion This study established an ensemble ML algorithm based on 3D-OCT images and clinical features for automatic detection of treatment responses to anti-VEGF treatment in DME patients to predict the efficacy of anti-VEGF treatment in DME patients and assist clinicians in optimal treatment decisions.

Keywords Diabetic macular edema, Anti-VEGF treatment, Machine learning, Optical coherence tomography.

†Yu Fang and Jianwei Lin contributed equally to this work.

*Correspondence:
Guihua Zhang
zgh@jsiec.org
[1]Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong, North Dongxia Road, Shantou 515041, Guangdong, China

Introduction

Diabetic macular edema (DME), a major complication of diabetic retinopathy (DR), is the leading cause of irreversible visual impairment and blindness in diabetic patients [1–3]. DR is an epidemic issue [4]. According to the International Diabetes Federation, approximately 537 million adults aged 20 to 79 were diagnosed with diabetes in 2021, with 103 million affected by DR [5]. By 2045, the number of DR patients is expected to reach 162 million, one-third of whom will develop DME [5–8].

At present, the most effective treatment for DME is intravitreal injection of anti-vascular endothelial growth factor (anti-VEGF) agents [9, 10]. Although anti-VEGF therapy showed better performances than retinal laser photocoagulation and steroids, approximately 30–40% of patients exhibit suboptimal or no response to the treatment [11–13]. The high cost and frequent injections in anti-VEGF therapy impose significant economic and adherence challenges to the patients [14, 15]. Together with the variability in responses to the anti-VEGF treatment it is warranted to develop an automated and accurate predictive system to predict the outcomes of anti-VEGF treatment.

Currently, DME diagnosis primarily depends on optical coherence tomography (OCT) images. However, the interpretation of the images is time-consuming and, even the experienced clinicians cannot fully explain the treatment outcomes of all patients [16].With the advancement in artificial intelligence, machine learning (ML) has been demonstrated the effectiveness and efficiency in handling image analysis in ophthalmology [17]. In 2021, a deep learning (DL) model has been developed for automated diagnosis and classification of DME based on the OCT images from various OCT devices [18]. In 2022, various classification and prediction models to forecasting the responses to anti-VEGF treatment in DME patients have been developed, and they exhibited outstanding performance in predicting structural and functional parameters and providing clinical recommendations up to six months in advance [19]. However, these models primarily focused on imaging features and did not incorporate clinical indicators. Previous reports indicated that systemic factors (gender, age, and hypertension) [20] and ocular structural changes (disorganization of the retinal inner layers (DRIL), hyper-reflective foci (HRF), external limiting membrane/ellipsoid zone (ELM/EZ)) can be the influential predictors for the anti-VEGF treatment responses [21–29]. Herein, in this study, we aimed to integrate the three-dimensional (3D)-OCT imaging features and clinical variables to construct ML models to predict the visual outcomes to the anti-VEGF treatment in DME patients.

Methods

This is a single-center cross-sectional study. The dataset was obtained from the patients diagnosed with DME and treated with anti-VEGF agents at the Joint Shantou International Eye Center (JSIEC) of Shantou University and the Chinese University of Hong Kong from January 2019 to January 2023. The anti-VEGF agents used included ranibizumab (Novartis Pharma Schweiz AG, Switzerland), aflibercept (Bayer AG, Germany) and conbercept (Chengdu Kanghong Biotech Co., Ltd., China). The dataset was divided into training set and test set in a ratio of 8:2. This study has been approved by the Human Medical Ethics Committee at JSIEC, which is in accordance with the Helsinki Declaration.

Data collection

The data collection included 3D-OCT images of the macular area before and after anti-VEGF treatment, and basic clinical information, including age, sex, eye laterality, best-corrected visual acuity (BCVA), examination time, diagnosis, diabetes stage, hypertension, lipid levels, and medications. The 3D-OCT images used in this study were obtained from Topcon DRIOCT Triton (Topcon Corporation, Japan), Topcon 3D-OCT 2000 (Topcon Corporation, Japan), and Cirus HD-OCT 5000 machines (Carl Zeiss Meditec, Germany). The imaging using Topcon DRI OCT Triton adopted the 3D Macula scanning mode with a 6×6 mm scan area centered on the macula, 512 A-scan and 256 B-scan signals, a scanning speed of 50,000 scans/second, a scanning depth of 2.3 mm, and a longitudinal resolution of 5–6 μm. The scanning mode for Topcon 3D OCT 2000 and Cirus HD-OCT 5000 instruments was 512×128, with an axial resolution of 5 rm. The 3D images centered on the fovea were generated after imaging. All 3D-OCT images collected in this study were captured by the experienced technicians.

The inclusion criteria of the study subjects included:

(1) Patients aged > 18 years old; (2) Diagnosed with diabetes mellitus and DME; (3) Received three consecutive injection of anti-VEGF agents within three months (with a 4–6-week interval), with baseline clinical parameters and 3D-OCT images before and at the final follow-up after the third injection, and the final follow-up time at one month and five days after the third injection; (4) Series of 3D-OCT images centered on the fovea, with no deviation from the foveal center, clear and complete retinal layer structures, no significant masking signals in the macular area, no other lesions or artifacts, and grayscale image channel selected; and (5) The 3D-OCT images with a quality score > 20 for Topcon DRIOCT Triton and Topcon 3D-OCT 2000 or signal value ≥ 4/10 for Cirrus HD-OCT 5000.

The exclusion criteria included:

(1) Concomitant other types of fundus diseases, including but not limited to Vogt-Koyanagi-Harada syndrome, retinal vein occlusion, Coats' disease, retinal artery occlusion, Eales disease, retinal detachment, and retinitis pigmentosa; (2) Anterior segment diseases affecting vision, such as keratitis, corneal opacities, and trachoma; (3) Received ocular surgery within six months before anti-VEGF treatment or during the study period, including cataract, glaucoma, vitreoretinal, and corneal surgeries; (4) Incomplete macular area structure or deviation of the image center from the foveal center in the 3D-OCT images; and (5) Despite passing the image quality assessment, various reasons resulted in missing sequences of two-dimensional tomography, leading to incomplete macular area structure or masking signals in the macular area. Image quality less than 20 for Topcon DRI OCT Triton and Topcon 3D OCT 2000, and image quality score less than 4/10 for Cirrus HD-OCT 5000.

Image quality classification: The 3D-OCT images were classified into good quality, qualified quality, and unacceptable quality. The images with unacceptable quality were excluded, while the images with good or qualified quality were included, with each set of images including three images before and after the anti-VEGF treatments at center fovea for annotation. The study initially involved 160 DME patients received three anti-VEGF treatments, resulting in a total of 190 eyes with 190 sets of 3D-OCT images. After quality screening, 142 patients with 170 eyes and 170 sets of 3D-OCT images were included.

Image preprocessing and annotation

The 3D-OCT images were obtained from three different OCT machines, each with varying dimensions. The images were processed using the Canny algorithm, Gaussian denoising, and pixel filling. The dark regions at the top and bottom of the images were cropped, retaining only the central region of interest. Finally, all images were cropped into a size of 384×384 pixels.

Two experienced readers independently annotated the 3D-OCT images from the Cirrus OCT reader or Topcon OCT reader. If their annotations were consistent, it became the final result. In cases of any inconsistency, the annotations were reviewed by a senior retinal specialist, and his annotation became the ultimate result. The annotations included the response classification, DME subtype, and definitions and annotations of biological markers on the OCT images (Table 1).

Variable analysis

To select the variables from the 21 parameters, we conducted the variable selection based on the correlation measures. For continuous variables, the Pearson correlation coefficient test was adopted. For categorical variables, Cramer's V test was applied. High correlation was defined as >0.9, and we excluded the variables with high correlations. The remaining variables were included as independent variables in the model construction.

Model training and evaluation

The 170 selected sets of 3D-OCT images were used to establish the anti-VEGF treatment response classification models using three commonly used convolutional neural network algorithms (Google Inception-v3, Xception, Inception-ResNet-V2). The dataset was split into training and test sets in an 8:2 ratio according to the number of patients. Each model generated a predicted probability of a good visual response. Eleven ML algorithms were used to construct the ML models based on the image prediction probabilities, corresponding clinical information, and annotated content. The selected ML algorithms included: logistic regression (LR), Naïve Bayes, support vector machine (SVM), K-Nearest neighbors (KNN), RandomForest, ExtraTrees, extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), GradientBoosting, adaptive boosting (AdaBoost), and multi-layer perceptron (MLP) [30–41]. The ML models were evaluated on the test set by the accuracy, sensitivity, specificity, F1 score, and area under the receiver operating characteristic (ROC) curve (AUC). To test for the accuracy of different ML models in anti-VEGF treatment response prediction and to select the ML models with outstanding performance, the ROC curve and the AUC were used as the comprehensive evaluation metrics.

Results

In total, the macular 3D-OCT images of 170 eyes from 142 DME patients were included. The mean age of the patients was 58.1 ± 10.9 years old, with 62 males (43.7%) and 80 females (56.3%). There were 81 left eyes (47.6%) and 89 right eyes (52.4%), with 87 sets of images from Topcon DRI OCT Triton, 59 sets from Topcon 3D OCT 2000, and 24 sets from Cirrus HD OCT 5000 (Table 2).

Based on the final BCVA improvement of logMAR >0.1, the 3D-OCT dataset was divided into good and poor visual response. We observed that there were 84 sets of 3D-OCT images with good visual response with, and 86 sets of 3D-OCT images with poor visual response. In the training set, there were 72 sets of 3D-OCT images classified as good visual response and 12 sets of 3D-OCT images classified as poor visual response. In the test set, there were 66 sets of 3D-OCT images classified as good visual response and 20 sets of 3D-OCT images classified as poor visual response (Supplementary Table 1).

For the 21 clinical and OCT features, high correlation was found for triglycerides, and it was not included for further analysis. The remaining 20 clinical and OCT features were included in the model construction.

Table 1 Annotations of OCT images and clinical features

Abbreviations	Full Terms	Classification Criteria
Annotations of OCT Images		
Eye	Eye laterality	0: Left; 1: Right
CFT1	Central foveal thickness of patients before treatment, μm	-
Cystic cavity	Maximum cystic cavity diameter / maximum macular thickness	0: ≤30%; 1: 30%~60%; 2: ≥60%
Maximum cystic cavity diameter	Maximum cystic cavity diameter, μm	-
ELM_EZ	External limiting membranes /ellipsoid zone in 1 mm diameter of fovea	0: The first and second outermost bands on OCT are visible; 1: Defect ≤ 200 μm; 2: Defect > 200 μm.
ELM_EZ integrity	Integrity of the external limiting membrane/ellipsoid zone within a 1 mm diameter of the fovea,, μm	-
DME Morphous	Diabetic macular edema morphous	1: Diffuse macular thickening; 2: Macular cystoid edema; 3: Serous retinal detachment; 4: More than two types.
DRIL	Disorganization of retinal inner layers	0: No; 1: Yes
HRF	Hyper-reflective foci	0: None; 1: 1 ~ 10; 2: More than 10
SF	Subretinal fluid	0: No; 1: Yes
Diagnosis	Staging of diabetic retinopathy	0: Non-Proliferative Diabetic Retinopathy; 1: Proliferative Diabetic Retinopathy
Clinical Features		
Sex	Patient's sex	0: Female; 1: Male
Age	Patient's age, years	-
logMar1	Vision of patients before treatment	-
Drugs	Medications used for intravitreal injection in patients	1: Conbercept; 2: Aflibercept; 3: Ranibizumab
Hypertension	Hypertension, mmHg	-
Cholesterol	Cholesterol, mmol/L	-
HDL	High-density lipoprotein, mmol/L	-
CR	Creatinine, ummol/L	-
LDT	Low-density lipoprotein, mmol/L	-
Triglycerides	Triglycerides, mmol/L	-

Table 2 Clinical data from 3D-OCT dataset

	Topcon DRI OCT Triton	Topcon 3D OCT 2000	Cirrus HD OCT 5000	Total Cohort
Patients, No.	70	52	20	142
Sex, No.(%)				
Male	30(48.40)	21(33.86)	11(17.74)	62
Female	40(50.00)	31(38.75)	9(11.25)	80
Study eye, No.(%)	87	59	24	170
Left	39(48.15)	28(34.57)	14(17.28)	81
Right	48(53.93%)	31(34.83)	10(11.23)	89
Age, mean (SD), y	60.68(8.96)	56.49(9.84)	52.46(15.82)	58.12(10.91)
3D-OCT image, No.	87	59	24	170

The performance analyses of each model in the training and test sets revealed that Gradient Boosting (accuracy of 0.912, AUC of 0.939, and F1 of 0.914), AdaBoost (accuracy of 0.853, AUC of 0.965, and F1 of 0.828), and LightGBM (accuracy of 0.794, AUC of 0.948, and F1 of 0.774) showed better performances in the test set (Supplementary Table 2). These three models were further selected for the construction of the ensemble algorithm,

which achieved an accuracy of 0.912, AUC of 0.976, and F1 score of 0.914 in the test set (Table 3). In the test set, the ensemble algorithm achieved a sensitivity of 0.941, specificity of 0.882, accuracy of 0.912, and AUC of 0.976 (0.937-1.000) (Fig. 1).

The results demonstrate that the model maintains high discriminative performance across different devices, particularly Topcon systems. The relatively lower AUC (0.67) observed with Zeiss devices may indicate performance variations, though further validation is needed due to the limited sample size.

For the ensemble algorithm construction, the contributing variables included the integrity of external limiting membranes/ellipsoid zone (importance score of 3.691), the vision before treatment (importance score of 3.452), subretinal fluid (importance score of 3.352), high-density lipoprotein-cholesterol (importance score of 1.668), medications used for intravitreal injections (importance score of 0.668), low-density lipoprotein-cholesterol (importance score of 0.449), hypertension importance score of 0.333), disorganization of retinal inner layers (importance score of 0.333), age (importance score of 0.035), creatinine (importance score of 0.006), maximum cystic cavity diameter (importance score of 0.005), and central foveal thickness before treatment (importance score of 0.005). The remaining variables had an importance score of 0. Using 1.5 times the mean as the relative threshold, variables with importance scores above this relative threshold were identified as effective variables, included the integrity of external limiting membranes/ellipsoid zone, the vision before treatment, subretinal fluid and high-density lipoprotein-cholesterol (Fig. 2).

Discussion

This study utilized 3D-OCT images from three different OCT machines to build a comprehensive database, integrating 20 variables of clinical parameters and OCT features, to develop an ensemble ML algorithm for anti-VEGF treatment response prediction in DME patients. In this study, after three consecutive anti-VEGF treatments, 84 eyes (49.4%) showed good visual response. The proportion of good visual response is consistent with previous studies [29]. Different response criteria selection may affect model performance. A previous study reported a model based on 10% decrease in total retinal thickness to predict the retinal thickness in DME patients after anti-VEGF treatment, achieving an AUC of 0.866 [41]. Moreover, an anatomical response classification model based on 10% decrease in central foveal thickness before and after treatment achieved an AUC of 0.923 [42]. In addition, a predictive model for anatomical responses in DME patients based on 25% decrease or 50 μm reduction of central macular thickness achieved an AUC of 0.923 and accuracy of 75% [43]. In this study, we adopted the final

Table 3 Performance of selected models and ensemble model

Model	ACC	AUC	95% CI	SEN	SPE	PPV	NPV	TPR	PPV	F1	Dataset
LightGBM	0.853	0.939	0.902–0.976	0.806	0.899	0.885	0.827	0.885	0.806	0.844	Train Set
LightGBM	0.794	0.948	0.881-1.000	0.706	0.882	0.857	0.750	0.857	0.706	0.774	Test Set
GradientBoosting	0.912	0.979	0.962–0.996	0.925	0.899	0.899	0.925	0.899	0.925	0.912	Train Set
GradientBoosting	0.912	0.939	0.853-1.000	0.941	0.882	0.889	0.937	0.889	0.941	0.914	Test Set
AdaBoost	0.875	0.948	0.916–0.980	0.940	0.812	0.829	0.933	0.829	0.940	0.881	Train Set
AdaBoost	0.853	0.965	0.921-1.000	0.706	1.000	1.000	0.773	1.000	0.706	0.828	Test Set
Model ensemble	0.912	0.978	0.960–0.996	0.925	0.899	0.899	0.925	0.899	0.925	0.912	Train Set
Model ensemble	0.912	0.976	0.937-1.000	0.941	0.882	0.889	0.937	0.889	0.941	0.914	Test Set

ACC: Accuracy; AUC: Area Under the Curve; 95% CI: 95% Confidence Interval; SEN: Sensitivity; SPE: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; TPR: True Positive Rate; F1: F1 Score; LightGBM: Light Gradient Boosting Machine; GradientBoosting: Gradient Boosting; AdaBoost: Adaptive Boosting; Model ensemble: Ensemble Model

To investigate the potential impact of imaging devices on model performance, we evaluated the prediction results separately for each device:

- Topcon 3D OCT-2000 ($n = 10$): AUC = 1.000
- Topcon DRI OCT Triton ($n = 20$): AUC = 1.000
- Zeiss Cirrus HD-OCT ($n = 4$): AUC = 0.667
- Test dataset (mixed) ($n = 34$): AUC = 0.976

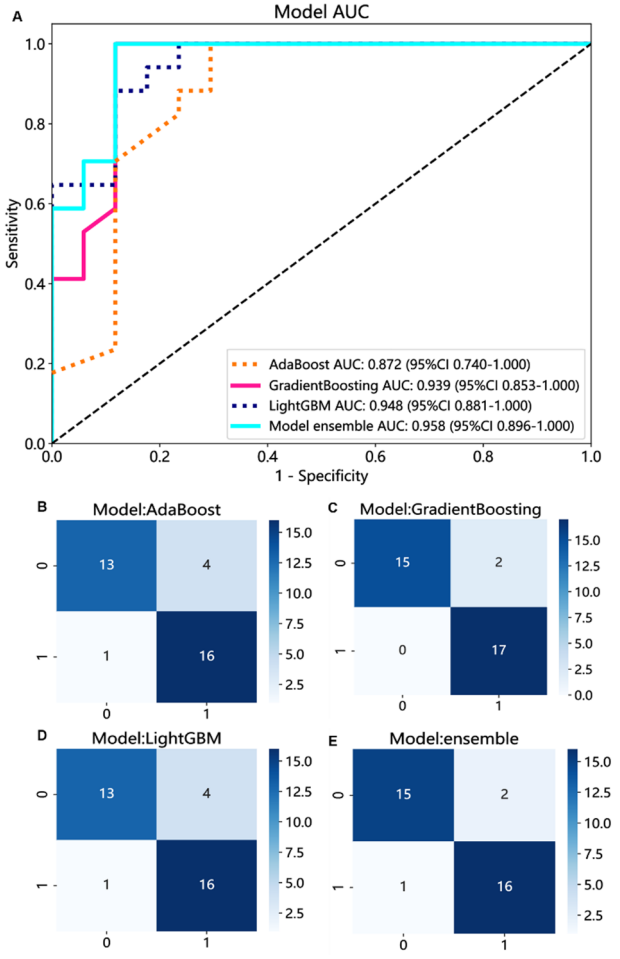

Fig. 1 Performance comparison between the ensemble model and individual constituent models

(**A**): ROC curves of each selected model and the ensemble model on the test set; (**B**): The confusion matrix of AdaBoost; (**C**): The confusion matrix of GradientBoosting; (**D**): The confusion matrix of LightGBM; (**E**): The confusion matrix of the ensemble model; AUC: Area Under the Curve; AdaBoost: Adaptive Boosting; GradientBoosting: Gradient Boosting; LightGBM: Light Gradient Boosting Machine; ensemble: Ensemble Model; 1: Good Vision Response; 0: Poor Vision Response

BCVA improvement of logMAR 0.1 as compared to the baseline BCVA as the classification criterion, achieved an AUC of 0.976 in the ensemble algorithm in the test dataset (Fig. 1).

Among the 11 tested ML models, LightGBM, GradientBoosting, and AdaBoost were applied in the ensemble model construction based on their performances. Consistently, a previous study developed an ensemble ML system comprising DL and CML models to predict central foveal thickness and BCVA after anti-VEGF treatment in DME patients with integration of information from multiple sources achieved an AUC of 0.94 for CFT prediction and 0.81 for BCVA prediction [44]. Other studies also showed that the ensemble models perform better than the individual constituent models [45, 46], consistent with our results that the ensemble model by LightGBM, GradientBoosting, and AdaBoost demonstrated better performance in classifying the responses to the anti-VEGF treatment in DME patients as compared to the individual ML models (Table 3).

In this study, four features, including the integrity of the external limiting membrane/ellipsoid zone, vision before treatment, subretinal fluid, and high-density lipoprotein-cholesterol, were identified as effective variables. In previous studies, the integrity of the external limiting membrane (ELM) and ellipsoid zone (EZ) has been widely recognized as an important indicator of the extent of retinal microstructural damage [47, 48]. Research has shown that structural disruption of the ELM and EZ is often closely associated with poor visual prognosis [49]. During anti-VEGF therapy, patients with intact ELM/EZ typically exhibit better treatment responses, suggesting that this region may play a critical role in photoreceptor functional recovery. Additionally, the concentration of high-density lipoprotein (HDL) is considered an independent risk marker associated with the severity of DR in patients with type 2 diabetes [50]. It is involved in the inflammatory activation of vascular walls and is linked to vascular endothelial dysfunction [51, 52]. Vision before treatment (baseline best-corrected visual acuity, BCVA) is also an important prognostic indicator. Studies by John et al. [53] and Ozlem et al. [54] demonstrated that patients with lower baseline BCVA tend to show poorer visual improvement after treatment. This may be because severely impaired vision suggests advanced structural damage to the macula, thereby limiting the potential for improvement. The presence of subretinal fluid (SRF) has been associated with favorable visual outcomes following anti-VEGF therapy. Jeff et al. [55] proposed that eyes with SRF exhibit higher VEGF levels and are therefore more likely to achieve significant anatomical and visual improvements after anti-VEGF treatment. Moreover, the resolution of SRF can serve as an indirect measure of treatment response. These clinical and imaging features not only emerged as effective variables in this study but also align with findings in the existing literature. This further validates their value as predictive factors in anti-VEGF therapy and enhances the interpretability of our predictive model.

There were several limitations in this study. First, the sample size was relatively small, and an external test set was not included, which may limit the generalizability of the findings. Second, although multiple studies have demonstrated varying treatment effects of different anti-VEGF agents in DME patients [56, 57], we did not conduct separate analyses based on the type of anti-VEGF agent used. Future studies could aim to develop algorithms that predict treatment outcomes for specific anti-VEGF agents. Third, although we normalized image

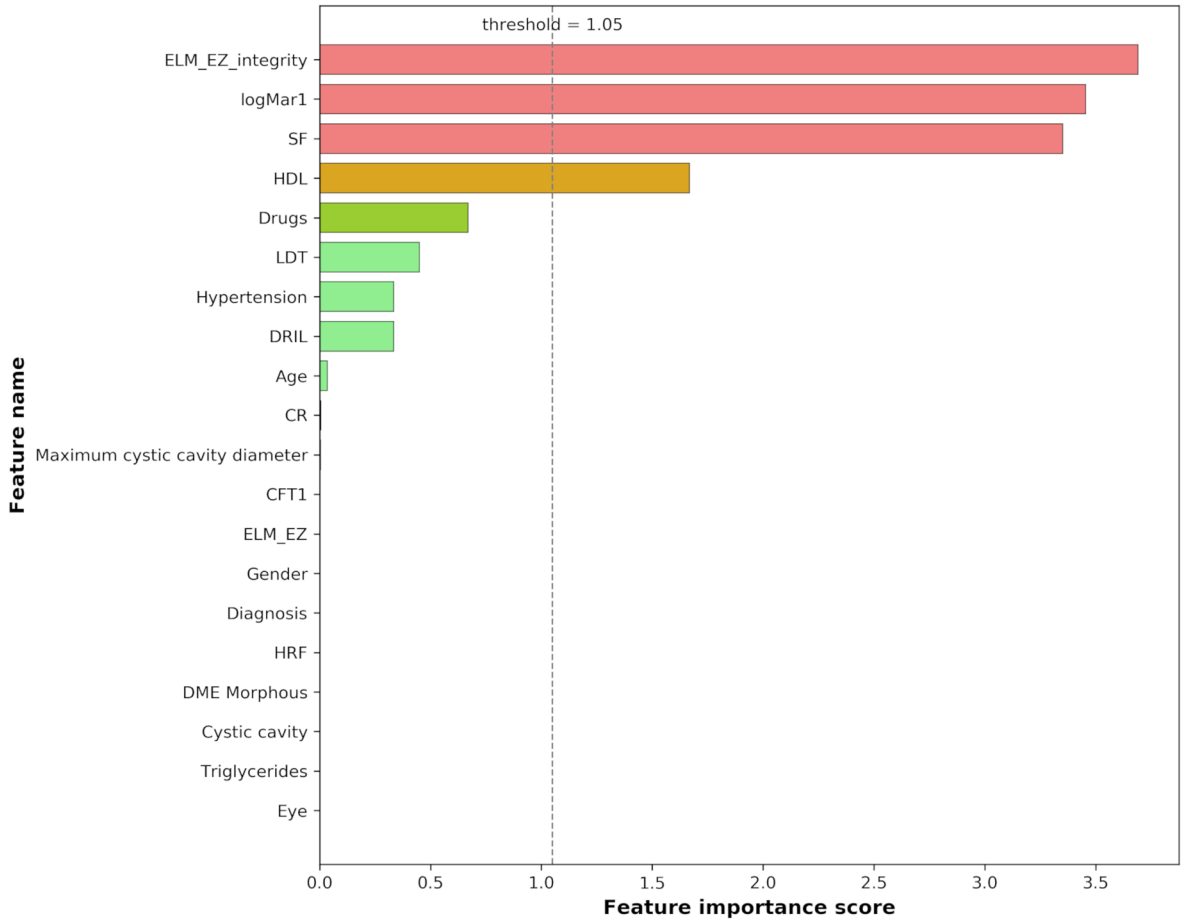

Fig. 2 Feature importance score of predictions from the ensemble model. ELM_EZ: external limiting membranes /ellipsoid zone in 1 mm diameter of fovea; logMar1: Vision of patients before treatment; SF: Subretinal fluid; HDL: High-density lipoprotein; Drugs: Medications used for intravitreal injection in patients; LDT: Low-density lipoprotein; DRIL: Disorganization of retinal inner layers; CR: Creatinine; CFT1: Central foveal thickness of patients before treatment; HRF: hyper-reflective foci; Threshold: 1.5 times the mean feature importance score

resolution and appearance during preprocessing, differences between imaging devices may still influence model behavior. Our stratified AUC analysis revealed that the model performed consistently well on Topcon 3D OCT-2000 and Triton devices, but its performance was lower on the limited subset of Zeiss Cirrus HD-OCT images ($n = 4$, AUC = 0.667). This discrepancy may reflect differences in image characteristics or clinical profiles associated with each device and also underscores the challenge posed by the small sample size. In future work, we plan to explicitly incorporate device type as a feature in the model or apply domain adaptation techniques to further mitigate such effects.

Conclusions

In conclusion, we developed an integrated ML system based on 3D-OCT images and clinical features for automatic detection of treatment responses to anti-VEGF treatment in DME patients to predict the efficacy of anti-VEGF treatment in DME patients and assist clinicians in optimal treatment decisions.

Abbreviations

DME	Diabetic macular edema
anti-VEGF	Anti-vascular endothelial growth factor
OCT	Optical coherence tomography
ML	Machine learning
DR	Diabetic retinopathy
DRIL	Disorganization of the retinal inner layers
HRF	Hyper-reflective foci
ELM/EZ	External limiting membrane/ellipsoid zone
JSIEC	The Joint Shantou International Eye Center
BCVA	Best-corrected visual acuity
LR	Logistic regression
SVM	Support vector machine
KNN	K-Nearest neighbors
XGBoost	Extreme gradient boosting
AdaBoost	Adaptive boosting
LightGBM	Light gradient boosting machine
MLP	Multi-layer perceptron
ROC	Receiver operating characteristic
AUC	Area under the receiver operating characteristic curve

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12886-025-04181-x.

Supplementary Material 1: Supplementary Table 1. 3D-OCT Subset Allocation Information. Supplementary Table 2. Performance of Each ML Model.

Acknowledgements
Not applicable.

Author contributions
YF and J-WL contributed equally to this work and share first authorship. YF and GZ proposed and designed the study. HZ and PX collected the data and/or images grading. J-WL analyzed the data. YF, TKN and GZ prepared the manuscript.

Funding
This work was supported by Natural Science Foundation of Guangdong Province (Grant No. 2025A1515010148) and Guangdong Provincial Science and Technology Special Fund (project code: 210719096902710). The funding bodies were not involved in the study, including the collection, analysis and interpretation of data and drafting the manuscript.

Data availability
The unpublished data are still being acquired for continuous screening programme. The data ownership belongs to the JSIEC. The authors are authorised to use the data from the JSIEC. The datasets used and/or analysed during the current study are available from the corresponding author [Guihua Zhang, fy@jsiec.org] on reasonable request.

Code availability

Project name
Ensemble Machine Learning Algorithm for Anti-VEGF Treatment Efficacy Prediction in Diabetic Macular Edema.

Project home page
No home page.

Operating system(s)
Ubuntu 18.04.6 LTS.

Programming Language
Python 3.5, C++.

Other requirements
Hardware: Intel E5-2620 V4 * 2, 64GB memory, 2 GeForce GTX 3090 GPUs.
Software: Intel MKL, Intel optimization for TensorFlow_cpu 1.12, CUDA 9.2, cuDNN 7.2.1, TensorFlow GPU version 1.12, PyTorch 1.12.1, MySQL Server (5.7.23), Anaconda 5.2.0.
Libraries/Frameworks: TensorFlow 4.0, Keras, OpenCV, NumPy, Scikit-learn, SciPy, Matplotlib, Pandas, Imgaug, SHAP.

License
MIT license.

Any restrictions to use by non-academics
Commercial use requires a separate license.

Declarations

Ethics approval and consent to participate
This study was approved by the Human Medical Ethics Committee of the Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong (EC20190911(4)-P14), which is in accordance with the Declaration of Helsinki. The Ethics Committee has approved the exemption of writing consent from patients as it was a retrospective study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Patient and public involvement
Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Received: 8 January 2025 / Accepted: 2 June 2025
Published online: 01 July 2025

References

1. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. Lancet. 2010;376:124–36.
2. Ferris FL, Patz A. Macular edema: a complication of diabetic retinopathy. Surv Ophthalmol. 1984;28:452–61.
3. Varma R, Bressler NM, Doan QV, et al. Prevalence of and risk factors for diabetic macular edema in the United States. JAMA Ophthalmol. 2014;132:1334–40.
4. Schmidt-Erfurth U, Garcia-Arumi J, Bandello F, et al. Guidelines for the management of diabetic macular edema by the European Society of Retina Specialists (EURETINA). Ophthalmologica. 2017;237:185–222.
5. Sun H, Saeedi P, Karuranga S, et al. IDF Diabetes Atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract. 2022;183:109119.
6. Oshitari T. Diabetic retinopathy: neurovascular disease requiring neuroprotective and regenerative therapies. Neural Regen Res. 2022;17:795–6.
7. Tang L, Xu GT, Zhang JF. Inflammation in diabetic retinopathy: possible roles in pathogenesis and potential implications for therapy. Neural Regen Res. 2023;18:976–82.
8. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. Ophthalmology. 2021;128:1580–91.
9. Ishibashi T, Li X, Koh A, et al. The REVEAL Study: ranibizumab monotherapy or combined with laser versus laser monotherapy in Asian patients with diabetic macular edema. Ophthalmology. 2015;122:1402–15.
10. Tatsumi T. Current treatments for diabetic macular edema. Int J Mol Sci. 2023;24:9591.
11. Bressler NM, Beaulieu WT, Glassman AR, et al. Diabetic Retinopathy Clinical Research Network. Persistent macular thickening following intravitreous aflibercept, bevacizumab, or ranibizumab for central-involved diabetic macular edema with vision impairment: a secondary analysis of a randomized clinical trial. JAMA Ophthalmol. 2018;136:257–69.
12. Virgili G, Parravano M, Menchini F, Evans JR. Anti-vascular endothelial growth factor for diabetic macular oedema. Cochrane Database Syst Rev. 2014;11:CD007419.
13. Nguyen QD, Brown DM, Marcus DM, RISE and RIDE Research Group, et al. Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. Ophthalmology. 2012;119:789–801.
14. Wells JA, Glassman AR, Ayala AR, et al. Diabetic Retinopathy Clinical Research Network. Aflibercept, bevacizumab, or ranibizumab for diabetic macular edema: two-year results from a comparative effectiveness randomized clinical trial. Ophthalmology. 2016;123:1351–9.
15. Ross EL, Hutton DW, Stein JD, et al. Diabetic Retinopathy Clinical Research Network. Cost-effectiveness of aflibercept, bevacizumab, and ranibizumab for diabetic macular edema treatment: analysis from the diabetic retinopathy clinical research network comparative effectiveness trial. JAMA Ophthalmol. 2016;134:888–96.
16. Xu F, Liu S, Xiang Y, et al. Prediction of the short-term therapeutic effect of anti-VEGF therapy for diabetic macular edema using a generative adversarial network with OCT images. J Clin Med. 2022;11:2878.
17. Lam C, Wong YL, Tang Z, et al. Performance of artificial intelligence in detecting diabetic macular edema from fundus photography and optical coherence tomography images: a systematic review and meta-analysis. Diabetes Care. 2024;47:304–19.
18. Tang F, Wang X, Ran AR, et al. A multitask deep-learning system to classify diabetic macular edema for different optical coherence tomography devices: a multicenter analysis. Diabetes Care. 2021;44:2078–88.

19. Xie H, Huang S, Liu Q, et al. The fundus structural and functional predictions of DME patients after anti-VEGF treatments. Front Endocrinol (Lausanne). 2022;13:865211.
20. Sophie R, Lu N, Campochiaro PA. Predictors of functional and anatomic outcomes in patients with diabetic macular edema treated with ranibizumab. Ophthalmology. 2015;122:1395–401.
21. Gerendas BS, Prager S, Deak G, et al. Predictive imaging biomarkers relevant for functional and anatomical outcomes during ranibizumab therapy of diabetic macular oedema. Br J Ophthalmol. 2018;102:195–203.
22. Shimura M, Yasuda K, Nakazawa T, et al. Visual outcome after intravitreal bevacizumab depends on the optical coherence tomographic patterns of patients with diffuse diabetic macular edema. Retina. 2013;33:740–7.
23. Sun JK, Lin MM, Lammer J, et al. Disorganization of the retinal inner layers as a predictor of visual acuity in eyes with center-involved diabetic macular edema. JAMA Ophthalmol. 2014;132:1309–16.
24. Radwan SH, Soliman AZ, Tokarev J, et al. Association of disorganization of retinal inner layers with vision after resolution of center-involved diabetic macular edema. JAMA Ophthalmol. 2015;133:820–5.
25. Kang JW, Chung H, Kim HC. Correlation of optical coherence tomographic hyperreflective foci with visual outcomes in different patterns of diabetic macular edema. Retina. 2016;36:1630–9.
26. Vujosevic S, Gatti V, Muraca A, et al. Hyperreflective retinal spots and visual function after anti-vascular endothelial growth factor treatment in center-involving diabetic macular edema. Retina. 2016;36:1298–308.
27. Maheshwary AS, Oster SF, Yuson RM, et al. The association between percent disruption of the photoreceptor inner segment-outer segment junction and visual acuity in diabetic macular edema. Am J Ophthalmol. 2010;150:63–e71.
28. Otani T, Yamaguchi Y, Kishi S. Correlation between visual acuity and foveal microstructural changes in diabetic macular edema. Retina. 2010;30:774–80.
29. Chen SC, Chiu HW, Chen CC, et al. A novel machine learning algorithm to automatically predict visual outcomes in intravitreal ranibizumab-treated patients with diabetic macular edema. J Clin Med. 2018;7:475.
30. Cox DR. The regression analysis of binary sequences. J R Stat Soc B. 1958;20(2):215–32.
31. John GH, Langley P. Estimating continuous distributions in Bayesian classifiers. Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI-95). 1995:338–345.
32. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–97.
33. Cover TM, Hart PE. Nearest neighbor pattern classification. IEEE Trans Inf Theory. 1967;13(1):21–7.
34. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
35. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn. 2006;63(1):3–42.
36. Chen T, Guestrin C, XGBoost:. A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016:785–794.
37. Ke G, Meng Q, Bai T, et al. LightGBM: A highly efficient gradient boosting decision tree. Adv Neural Inform Process Syst (NeurIPS). 2017;30:3146–54.
38. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann Stat. 2001;29(5):1189–232.
39. Freund Y, Schapire RE. Experiments with a new boosting algorithm. Proceedings of the Thirteenth International Conference on Machine Learning (ICML-96). 1996:148–156.
40. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. Psychol Rev. 1958;65(6):386–408.
41. Rasti R, Rabbani H, Mehridehnavi A, et al. Deep learning-based single-shot prediction of differential effects of anti-VEGF treatment in patients with diabetic macular edema. Biomed Opt Express. 2020;11:1139–52.
42. Alryalat SA, Al-Antary M, Arafa Y, et al. Deep learning prediction of response to anti-VEGF among diabetic macular edema patients: Treatment Response Analyzer System (TRAS). Diagnostics (Basel). 2022;12:312.
43. Cao J, You K, Jin K, et al. Prediction of response to anti-vascular endothelial growth factor treatment in diabetic macular oedema using an optical coherence tomography-based machine learning method. Acta Ophthalmol. 2021;99:e19–27.
44. Liu B, Zhang B, Hu Y, et al. Automatic prediction of treatment outcomes in patients with diabetic macular edema using ensemble machine learning. Ann Transl Med. 2021;9:43.
45. Orlando JI, Prokofyeva E, Del Fresno M, et al. An ensemble deep learning-based approach for red lesion detection in fundus images. Comput Methods Programs Biomed. 2018;153:115–27.
46. Orlando JI, Gerendas BS, Riedl S, et al. Automated quantification of photoreceptor alteration in macular disease using optical coherence tomography and deep learning. Sci Rep. 2020;10:5619.
47. Maheshwary AS, Oster SF, Yuson RM, Cheng L, Mojana F, Freeman WR. The association between percent disruption of the photoreceptor inner segment-outer segment junction and visual acuity in diabetic macular edema. Am J Ophthalmol. 2010;150(1):63–e671.
48. Saxena S, Sadda SR. Focus on external limiting membrane and ellipsoid zone in diabetic macular edema. Indian J Ophthalmol. 2021;69(11):2925–7.
49. Kessler LJ, Auffarth GU, Bagautdinov D, Khoramnia R. Ellipsoid Zone Integrity and Visual Acuity Changes during Diabetic Macular Edema Therapy: A Longitudinal Study. J Diabetes Res. 2021;2021:8117650.
50. Dimitriou E, Sergentanis TN, Lambadiari V, Theodossiadis G, Theodossiadis P, Chatziralli I. Correlation between Imaging Morphological Findings and Laboratory Biomarkers in Patients with Diabetic Macular Edema. J Diabetes Res. 2021;2021:6426003.
51. Sotiriou SN, Orlova VV, Al-Fakhri N, et al. Lipoprotein(a) in atherosclerotic plaques recruits inflammatory cells through interaction with Mac-1 integrin. FASEB J. 2006;20(3):559–61.
52. Wu HD, Berglund L, Dimayuga C, et al. High lipoprotein(a) levels and small apolipoprotein(a) sizes are associated with endothelial dysfunction in a multiethnic cohort. J Am Coll Cardiol. 2004;43(10):1828–33.
53. Wells JA, Glassman AR, Ayala AR, et al. Aflibercept, Bevacizumab, or Ranibizumab for Diabetic Macular Edema: Two-Year Results from a Comparative Effectiveness Randomized Clinical Trial. Ophthalmology. 2016;123(6):1351–9.
54. Eski Yucel O, Birinci H, Sullu Y. Outcome and Predictors for 2-Year Visual Acuity in Eyes with Diabetic Macular Edema Treated with Ranibizumab. J Ocul Pharmacol Ther. 2019;35(4):229–34.
55. Park J, Felfeli T, Kherani IZ, Altomare F, Chow DR, Wong DT. Prevalence and clinical implications of subretinal fluid in retinal diseases: a real-world cohort study. BMJ Open Ophthalmol. 2023;8(1):e001214.
56. Diabetic Retinopathy Clinical Research Network, Wells JA, Glassman AR, Ayala AR, et al. Aflibercept, bevacizumab, or ranibizumab for diabetic macular edema. N Engl J Med. 2015;372:1193–203.
57. Virgili G, Parravano M, Evans JR, et al. Anti-vascular endothelial growth factor for diabetic macular oedema: a network meta-analysis. Cochrane Database Syst Rev. 2017;6:CD007419.

Publisher's note