



OPEN

DATA DESCRIPTOR

GRAPE: A multi-modal dataset of longitudinal follow-up visual field and fundus images for glaucoma management

Xiaoling Huang¹, Xiangyin Kong², Ziyang Shen³, Jing Ouyang², Yunxiang Li⁴, Kai Jin¹✉ & Juan Ye¹✉

As one of the leading causes of irreversible blindness worldwide, glaucoma is characterized by structural damage and functional loss. Glaucoma patients often have a long follow-up and prognosis prediction is an important part in treatment. However, existing public glaucoma datasets are almost cross-sectional, concentrating on segmentation on optic disc (OD) and glaucoma diagnosis. With the development of artificial intelligence (AI), the deep learning model can already provide accurate prediction of future visual field (VF) and its progression with the support of longitudinal datasets. Here, we proposed a public longitudinal glaucoma real-world appraisal progression ensemble (GRAPE) dataset. The GRAPE dataset contains 1115 follow-up records from 263 eyes, with VFs, fundus images, OCT measurements and clinical information, and OD segmentation and VF progression are annotated. Two baseline models demonstrated the feasibility in prediction of VF and its progression. This dataset will advance AI research in glaucoma management.

Background & Summary

Glaucoma is a degenerative optic neuropathy characterized by the loss of retinal ganglion cells that frequently causes irreversible blindness^{1,2}. It has affected more than 64.3 million individuals worldwide and the number may be increased to 111.8 million in 2040³. It has different subtypes and a wide range of inconspicuous risk factors, such as the optic nerve head (ONH) damage, visual field (VF) loss and intraocular pressure (IOP) elevation⁴. Timely detection and treatment can prevent glaucoma from getting worse⁵.

Due to the difficulty in treatment, the glaucoma patients are often accompanied with a chronic progressive course⁶. Regular follow-ups with VF reviewing by standard automated perimetry (SAP) and IOP monitoring show great significance to the visual function evaluation of glaucoma patients⁷. The color fundus photograph (CFP) and optical coherence tomography (OCT) are implemented to detect structural damage. The routine examinations during follow-up of one glaucoma patient are shown in Fig. 1. At the first time, relatively complete examinations would be applied of one suspected glaucoma patient, for the doctor to generate the preliminary diagnosis. Then the patient with intraocular hypertension would take the medicine to decrease IOP to prevent further deterioration of visual function, and visit regularly to check IOP and VF, sometimes fundus structure if necessary. If IOP is uncontrolled, additional medical treatments, laser treatments and surgical treatments would be performed.

With the development of artificial intelligence (AI), several researches have paid attention to glaucoma management, and used several biometric parameters or medical images to predict the onset and progression of glaucoma^{8–12}. On account of the considerable test-retest variability and strong patient cooperation requirement of VF tests^{13–15}, previous studies have also proposed various machine learning (ML) and deep learning (DL) models to predict VF from structural parameters or images, with tolerable estimation error^{16–19}. The development of these AI models about glaucoma assessment needs complete follow-up examination data of patients.

¹Eye Center, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, 310003, China. ²State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, 310013, China. ³Zhejiang Baima Lake Laboratory Co., Ltd, Hangzhou, 310051, China. ⁴Department of Radiation Oncology, UT Southwestern Medical Center, Dallas, TX, 75235, USA. ✉e-mail: jinkai@zju.edu.cn; yejuan@zju.edu.cn

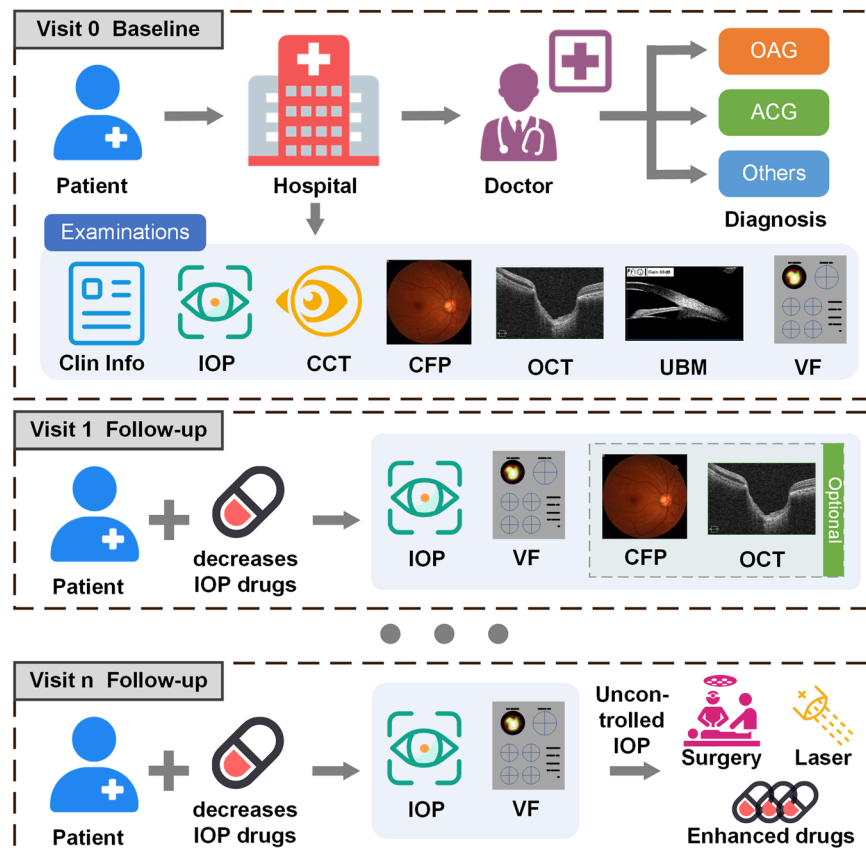


Fig. 1 The routine medical process of glaucoma patients. The complete examinations tend to be accomplished to confirm the diagnosis at the first time of one patient. Subsequently, the patient with decreasing IOP drugs would do regular follow-up to check IOP, VF and other examinations. If IOP is uncontrolled, additional medical treatments, laser treatments and surgical treatments are necessary.

Many well-known glaucoma public datasets of retinal images have been established for automated glaucoma screening, including RIGA²⁰, ORIGA²¹, DRISHTI-GS²², REFUGE²³ and PAPILA²⁴. These datasets mainly concentrated on the segmentation of optic disc (OD) and optic cup (OC), emphasizing the discrepancy on fundus characteristics between glaucoma and normal patients. However, there are few public datasets about glaucoma progression research, and some are listed below:

- UWHVF²⁵: This is an open-source VF dataset from the University of Washington with 28943 VFs. It contains longitudinal VFs from one eye or both eyes of one patient. Pointwise sensitivities were extracted from Humphrey Field Analyzer (HFA) with other related VF parameters calculated from it. The progression analysis was performed on part of VFs.
- Annotated Glaucoma Medication Dataset²⁶: This is an open-source dataset including 480 clinical notes from the Electronic Health Record (EHR) system. All progress notes were annotated for glaucoma medication name, route, frequency, dosage, and drug use. It is mainly used for natural language processing model development.

In general, most datasets containing CFP or OCT images focused on glaucoma diagnosis, that longitudinal changes generated during follow-up cannot be evaluated. While the datasets for progression detection only include VF values or medical records, without comprehensive reflecting the condition of glaucoma patients. Future VF progression prediction from IOP²⁷ or CFP¹¹ cannot be realized on these datasets.

In this case, we proposed a glaucoma real-world appraisal progression ensemble (GRAPE) dataset, consisting of the clinical data, VF values, OCT measurements and CFPs at the baseline and during the follow-up visits. It records the full follow-up procedure of glaucoma patients. This dataset could be used for progress prediction and VF estimation of glaucoma. It will promote AI-based researches in glaucoma management.

Methods

Patient inclusion. All data in the GRAPE dataset was collected in the Eye Center at the Second Affiliated Hospital of Zhejiang University (ZJU), that contains 1115 records of 263 eyes from 144 glaucoma patients from 2015 to 2022, with ages ranging from 18 to 81 years. We excluded the patients under 18 years old. Written informed consent complying with the requirement of the Medical Ethics Committee of ZJU was signed by

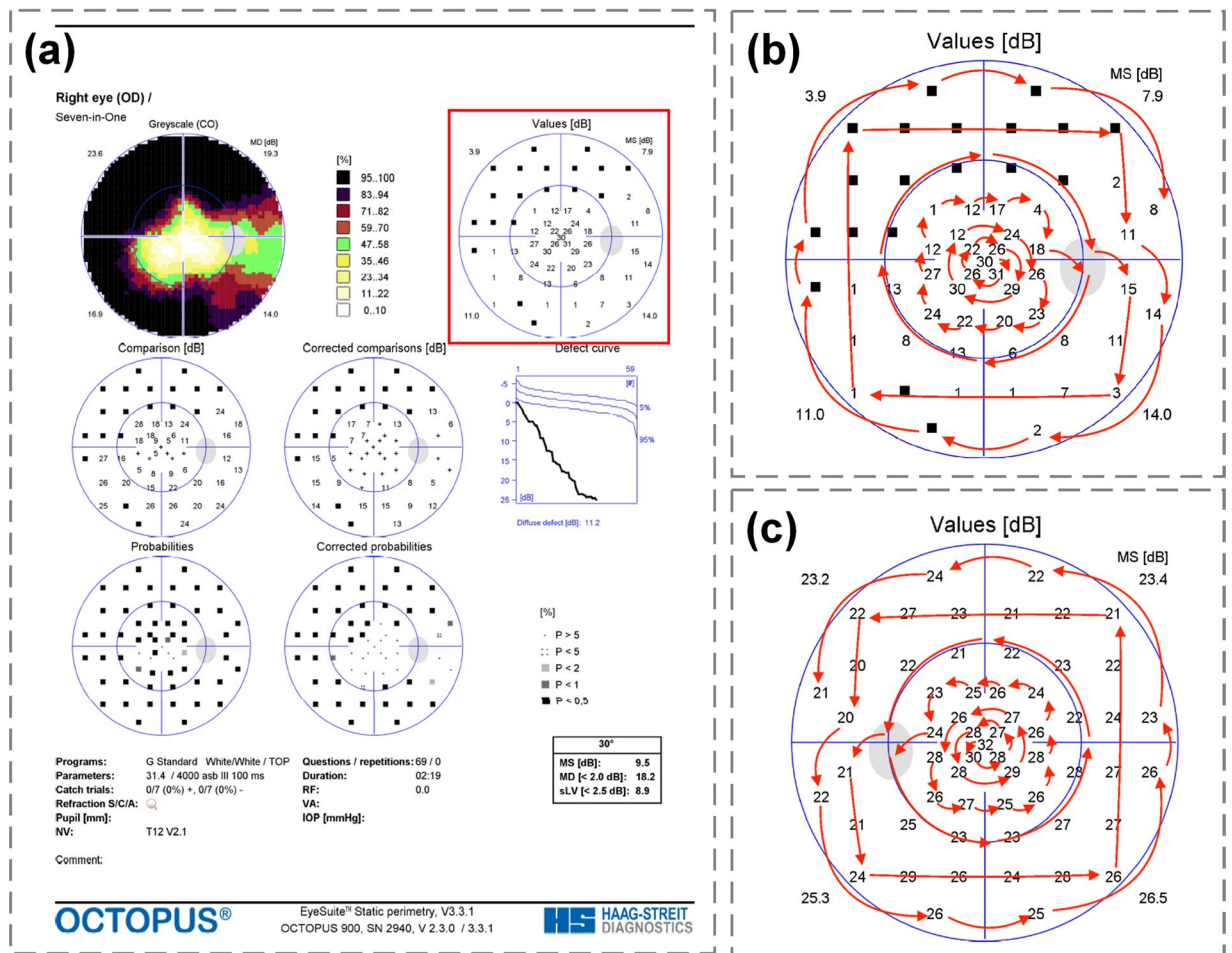


Fig. 2 The visual field (VF) report and the extraction order of the values. (a) an example of a VF report from Octopus perimeter. We used the light sensitivity values in the “Values” in the red box in the GRAPE dataset. (b) The extraction order of VF values of the right eye. (c) The extraction order of VF values of the left eye.

every participant. This study is a registered clinical study (A New Technique for Retinal Disease Treatment, ClinicalTrials.gov identifier: NCT04718532). Ethical approval for the study was obtained from Ethics Committee of ZJU-2 (No Y2020–1027). The research adhered to the tenets of the Declaration of Helsinki and the Health Portability and Accessibility Act.

This study only included patients with a definite diagnosis of both open-angle glaucoma (OAG) and angle-closure glaucoma (ACG). The specialists comprehensively considered the glaucoma examination of patients, including medical history, function assessment (VF), structural assessment (CFP or OCT or both), IOP measurement, CCT measurement, anterior chamber angle evaluation. A suspected glaucoma patient would have further visits to confirm the diagnosis. These glaucoma patients were mainly diagnosed by senior outpatient clinicians, and subsequently confirmed by 3 glaucoma specialists to be included in this dataset, according to the American Academy of Ophthalmology’s Preferred Practice Pattern of glaucoma guidelines^{28,29}. The exclusion criteria are: (1) patients with other optic nerve diseases; (2) patients with severe retinal diseases (e.g. vitreous hemorrhage and retinal detachment); (3) patients with severe dioptric media turbidity; (4) patients who had undergone glaucoma surgery.

Data collection. The collection process of different data categories is listed below:

VF. The VFs included in the GRAPE dataset were measured by two experienced technicians using the G1 program test pattern with stimulus size III by OCTOPUS 900 perimeter (HAAG-STREIT, Switzerland). The VFs with false-negative rate $\geq 30\%$ or false-positive rate $\geq 30\%$ were identified as unreliable examinations and were excluded^{30,31}. The VF values were extracted from the “Values” on VF reports, as shown in Fig. 2a. The “Values” in Octopus VF reports is equal to the light sensitivity values in VF reports measured by Humphrey Field Analyzer (HFA). The VFs were extracted with certain order that is mirror for left and right eyes to ensure all the VFs with the same format, as right eye shown in Fig. 2b and left eye shown in Fig. 2c. The black square, “■”, indicates that this eye is not able to perceive the lowest intensity of light at this location, and we assigned value -1 for it in the dataset. Likewise, two blind points were assigned value -1 . The VFs in our dataset were finally presented as digital format in the table sheet.

CFP. The CFPs were obtained by two experienced technicians using a 50° field of view centered at macula by the TRC-NW8 Fundus Camera (TOPCON, Japan), CR-2 PLUS AF Digital Retinal Camera (CANON, Japan) and CR-2 AF Digital Retinal Camera (CANON, Japan) without pupil dilation. The pictures were saved in JPG format with 3 resolutions of 1556 × 1556 pixels, 2136 × 2136 pixels, and 1611 × 1611 pixels for these three cameras, respectively. We excluded the poor-quality images that 50% of the area is obscured or only part of OD is visible. And we included the image with the best quality if there were more than one image of each eye at the same visit. The image preprocessing of extracting the region of interest (ROI) and annotating the OD segmentation was described below.

Retinal nerve fiber layer (RNFL) thickness. The RNFL thickness measurements at baseline in the GRAPE dataset were acquired by two experienced technicians using CIRRUS HD-OCT 5000 (Carl Zeiss Meditec, United States of America). The ONH and RNFL Analysis pattern for Optic Disc 200 × 200 was used to detect and calculate the total and 4 sectoral (superior, nasal, inferior and temporal) mean peripapillary RNFL thickness.

IOP. The values of IOP were measured by two technicians using non-contact tonometer NT-530P (NIDEK, Japan). The patients were informed that there would be air ejection during the measurement in advance, to avoid the error caused by the scare for the sudden air ejection. Each eye was automatically measured for 3 times and the average value was used. The glaucoma patients included used the IOP control medications in the follow-up period.

Central corneal thickness (CCT). The values of CCT were measured by two technicians using non-contact tonometer NT-530P (NIDEK, Japan), same to the IOP. The patients were with the correct sitting position and focused on the indicator light, then technicians could measure the value of CCT.

Basic clinical data. These clinical data was acquired from the EHR system. The basic data and medical history were inquired by clinicians or nurses during the attendance process.

Data summary. The data provided in the GRAPE dataset and the format are listed in Table 1. There are two sheets in the table for the baseline data and the follow-up data, respectively. The baseline visit is defined as the first time the patient received a comprehensive evaluation of glaucoma, including medical history, function assessment (VF), structural assessment (CFP, OCT or both), IOP measurement, CCT measurement, anterior chamber angle evaluation. If the patient had prior visit record for other ocular diseases or only with VF examination, this visit record would not be included in the dataset.

There are 263 multi-modal records in the baseline data, including subject number, laterality, age, gender, IOP, CCT, total visits, progression status, category of glaucoma, OCT RNFL thickness, corresponding CFP, acquisition device, resolution and VF. And there are 1115 visit records in follow-up data, consisting of subject number, laterality, visit number, interval years, IOP, corresponding CFP, acquisition device, resolution and VF. The quantification methods of progression are introduced below. Besides, the original CFPs were placed in one folder, while the processed images and the annotation files were placed in another folder. The image preprocessing method is also introduced below.

Image preprocessing. The fundus manifestation of glaucoma contains OD rim narrowing, cup-to-disc ratio (CDR) increasing, large extent of parapapillary atrophy, and RNFL defect^{28,29}, mainly focused on the area around OD. For AI-based model to identify the characteristics more facilely, ROI, a square area around OD, was clipped by one experienced ophthalmologist. Generally, the OD was regarded as the center of the clipped image. Besides, in view of the large extent of parapapillary atrophy is one of the major features, we preserved this structure in the ROI as possible. The ROIs of CFPs were cropped into 453 × 453 pixels, 566 × 566 pixels and 432 × 432 pixels for three devices, respectively. The whole procedure of image disposing is shown in Fig. 3. In addition, we hid all personal information on CFPs.

To singularize the feature of CDR increasing and OD rim narrowing, OD and OC segmentation on ROI images was performed by one ophthalmologist using the label tool. The red circle is the border of OC and green of OD in the annotated images. We also provided the original annotation files, JSON files, that record the XY coordinates, and the codes that could draw the label line on the ROI images.

Quantification of VF progression. The Comparison Values, that are the total deviation (TD) values on HFA reports, were calculated by light sensitivity values and normative values. The normative values were obtained from the device according to the different ages. TD values were calculated for all 59 locations excluding the two blind spots, with the locations same to that of light sensitivity values. Because the TD values less than 5 were displayed as “+” in Comparison of Octopus VFs, we calculated the TD values according to the original sensitivity values and age, conforming to the definition provided by the Imaging and Perimetry Society³².

There are several different methods for detecting glaucomatous VF progression^{33,34}. We chose three commonly used automated progression criteria, two based on point-wise linear regression (PLR) analysis and one based on mean deviation (MD) slope, on account of the fitness for Octopus VFs. For PLR, the baseline VF of one eye is defined as progressing if the number of test points with a significant negative regression slope ($P < 0.01$) greater than or equal to two (PLR2) or three (PLR3). For MD slope, when the slope of MD over follow-up time is presented with negative linear regression and P value is less than 0.05, this eye is regarded as progression. We calculated the progression by the visual Felds package³⁵, an open-source software to analyze VF based on R environment. This package has been modified to be compatible with Octopus VF.

| | Data | Format | Detail description |
|----------------------|--------------------|-------------------|--|
| Baseline | Age | Number | The age at the first visit |
| | Gender | F/M | “F” for female and “M” for male |
| | CCT | Number | The value of CCT |
| | Times of visits | Number | Times of visits during follow-up |
| | Progression-PLR2 | 0/1 | “0” for no progression, and “1” for progression |
| | Progression-PLR3 | 0/1 | “0” for no progression, and “1” for progression |
| | Progression-MD | 0/1 | “0” for no progression, and “1” for progression |
| | Category | OAG/ACG | The two categories of glaucoma |
| | OCT RNFL thickness | A group of number | Including mean (“Mean”), superior (“S”), nasal (“N”), inferior (“I”) and temporal (“T”) peripapillary RNFL thickness |
| Baseline & follow-up | Subject number | Number | The serial number of inclusion subjects in this dataset |
| | Laterality | OD/OS | The laterality of one eye, “OD” for right eye and “OS” for left eye |
| | Visit number | Number | The times for follow-up visits from baseline |
| | Interval year | Number | The interval years between each visit and baseline |
| | IOP | Number | The value of IOP |
| | VF | A group of number | Including the Octopus VF values of 61 points |
| | CFP | Image | Including the original CFP, ROI and ROI with OD/OC annotation |
| | Acquisition Device | Device name | The devices taking the CFPs, including 3 types of devices |
| | Resolution | Number × number | The resolution of CFP, corresponding to the Acquisition Device, including 3 types, 1556 × 1556, 2136 × 2136, 1611 × 1611 |

Table 1. Summarization and detail description of the GRAPE dataset.

Data Records

The GRAPE dataset is available at Figshare³⁶. The main part of the dataset includes 6 parts, 1 Excel file, 4 zipped files and 1 python file. The zipped file could be decompressed to the corresponding folders.

The Excel file, “VF and clinical information”, contains 2 sheets, titled “Baseline” and “Follow-up”. There are 263 records representing 263 eyes and 80 columns in sheet “Baseline”, the last 61 columns of which are “VF”. These columns are in turn: “Subject Number” (inclusion subjects numbered as 1–144), “Laterality” (OD for right eye and OS for left eye), “Age”, “Gender” (M for male and F for female), “IOP”, “CCT”, “Total Visits” (the total times of visits for each eye), “Progression Status-PLR2” (progression status defined by PLR2 method, 0 for non-progression and 1 for progression), “Progression Status-PLR3” (progression status defined by PLR3 method, 0 for non-progression and 1 for progression), “Progression Status- MD” (progression status defined by MD method, 0 for non-progression and 1 for progression), “Category of Glaucoma” (OAG for open-angle glaucoma and ACG for angle-closure glaucoma), “OCT RNFL thickness” (“Mean” for total mean RNFL peripapillary thickness, “S” for superior, “N” for nasal, “I” for inferior, and “T” for temporal), “Corresponding CFP” (named as “subject number_laterality_visit number. jpg”), “Acquisition Device” (the device that taking the corresponding CFP), “Resolution” (the resolution of the corresponding CFP) and “VF” (the last 61 columns). Besides, there are 1115 records representing 1115 visits and 69 columns in sheet “Follow-up”, the last 61 columns of which are “VF” as well. These columns are in turn: “Subject Number”, “Laterality”, “Visit Number” (the times for follow-up visit from baseline), “Interval Years” (the interval years for follow-up visit from baseline), “IOP”, “Corresponding CFP”, “Acquisition Device”, “Resolution” and “VF”. The symbol “/” in the sheet represents that no corresponding CFP or OCT examination was generated in this visit. The detailed description of the dataset is shown in Table 1.

The folder “CFPs” contains 631 original CFPs from non-mydriatic digital camera, with personal information hidden. These images are named as “subject number_laterality_visit number. jpg”, corresponding to the records in Excel file.

The folder “ROI images” contains 631 ROI of CFPs cropped from the original images, and the naming of them is consistent with that of original images.

The folder “Annotated images” includes 631 ROI images with OD/OC segmentation. The green contour is for OD and red for OC, respectively. And the naming is consistent with the original image and ROI, as well.

The folder “json” contains 631 original annotation files in JSON format, that record the XY coordinates. The naming method is consistent with the images, in the format “subject number_laterality_visit number. json”. The JSON file can be opened using Notebook and some label tool software.

The python file, “draw.py”, records the codes for contours drawing. Running the codes in the “draw.py” can annotate the segmentation in json files on ROI of CFPs, and the results will be output in the folder “Annotated Images”.

Technical Validation

Dataset characteristics. There are 1115 records of 263 eyes from 144 glaucoma patients in the GRAPE dataset. Each visit of the patient serves as a record, and each eye has 3–9 times of visits. The interval between adjacent visit is more than 5 months. The mean age is 42.49 years old. Both genders and eye laterality remain a certain balance. The majority of patients in this dataset are OAG, may due to the requirement of complete examination for diagnosis of OAG and its long follow-up time for management. The detailed data characteristics are shown in Table 2.

| Item | Value |
|--------------------------------------|---------------|
| The number of patients | 144 |
| The number of eyes | 263 |
| The number of records | 1115 |
| Average age (SD) | 42.51 (15.41) |
| Male (%) | 75 (52.08%) |
| Right eye (%) | 130 (49.43%) |
| OAG (%) | 139 (96.53%) |
| Baseline MD (dB) (SD) | -7.11 (5.71) |
| Total MD (dB) (SD) | -7.21 (5.71) |
| Average follow-up times (SD) | 4.21 (1.31) |
| Average follow-up years (SD) | 2.51 (1.01) |
| Average interval years (SD) | 0.81 (0.41) |
| The number of progression - PLR2 (%) | 40 (15.21%) |
| The number of progression - PLR3 (%) | 14 (5.32%) |
| The number of progression - MD (%) | 27 (10.27%) |

Table 2. The data characteristics of the GRAPE dataset.

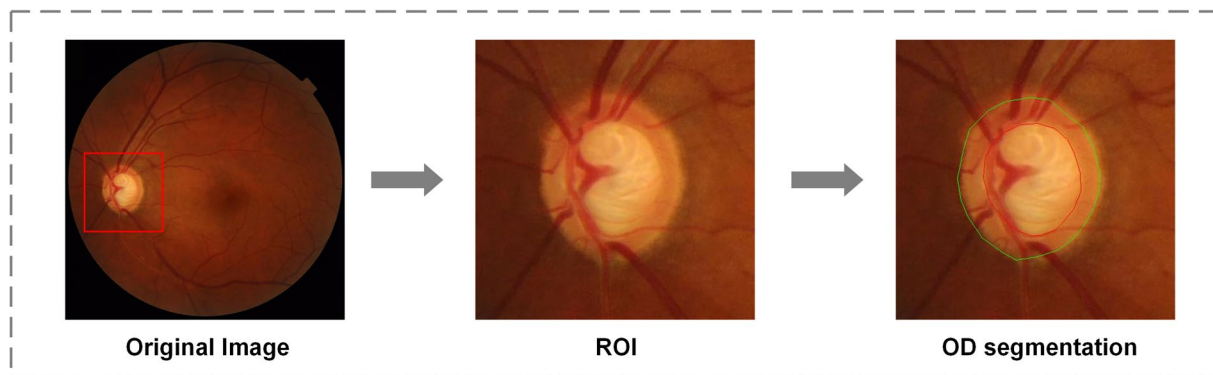


Fig. 3 The preprocessing procedure of color fundus photographs (CFPs). The region of interest (ROI) was clipped from the original image, and the annotation of optic disc (green line) and optic cup (red line) was labeled by experienced ophthalmologist on ROI.

There are some statistic characteristics for VFs as well. Average light sensitivities of all follow-up VFs are shown in Fig. 4a. The original light sensitivities of central VF are highest and tend to be the last to be impaired. The nasal VF values are lowest due to the damage usually starting from the nasal side. Average progression slopes are shown in Fig. 4b. The negative progression is concentrated in the lower part of the paracentral field and temporal area. We suspect that some patients had been presented with impaired lateral VF, and the paracentral VFs were damaged subsequently.

Baseline model validation for VF progression prediction. Because of the longitudinal and multi-modal data during the entire follow-up included, the GRAPE dataset is available for multi-purpose models training. It is important for patients that could be aware of their future condition at the first visit and it could also guide management. The prediction of whether VF progression or not has been realized in previous research from CFPs⁴¹ based on deep convolutional neural networks (CNNs). However, more experiments are needed to valid the prediction. In this part, we applied a common baseline model, Resnet-50, to show the technical validation of our dataset.

The CFPs of each eye at the baseline was considered as input, while the progression was defined by PLR2, PLR3 and MD slope as output, that each eye was classified as “progression” or “non-progression”.

The 263 CFPs were resized to $224 \times 224 \times 3$ and the mean variance standardization method was used for normalization. We trained the basic CNN, Resnet-50, using cross entropy loss function for 264 epochs with a batch size of 4, with Adam optimizer. In addition, we used annealed cosine to update the learning rate from 10^{-4} to 10^{-5} . The Receiver Operating Characteristic Curves were used to evaluate the performance of ResNet-50.

We evaluated the results of predicting performance by the area under receiver operating characteristic curve (AUC), shown in Fig. 5. The AUC of the models of PLR2, PLR3 and MD is 0.71, 0.80 and 0.73. Besides, the accuracy of the models of PLR2, PLR3 and MD is 0.75, 0.91 and 0.81, respectively. It demonstrated that the GRAPE dataset could be used for VF progression prediction and the model with data partitioning defined by PLR3 has the best classification performance.

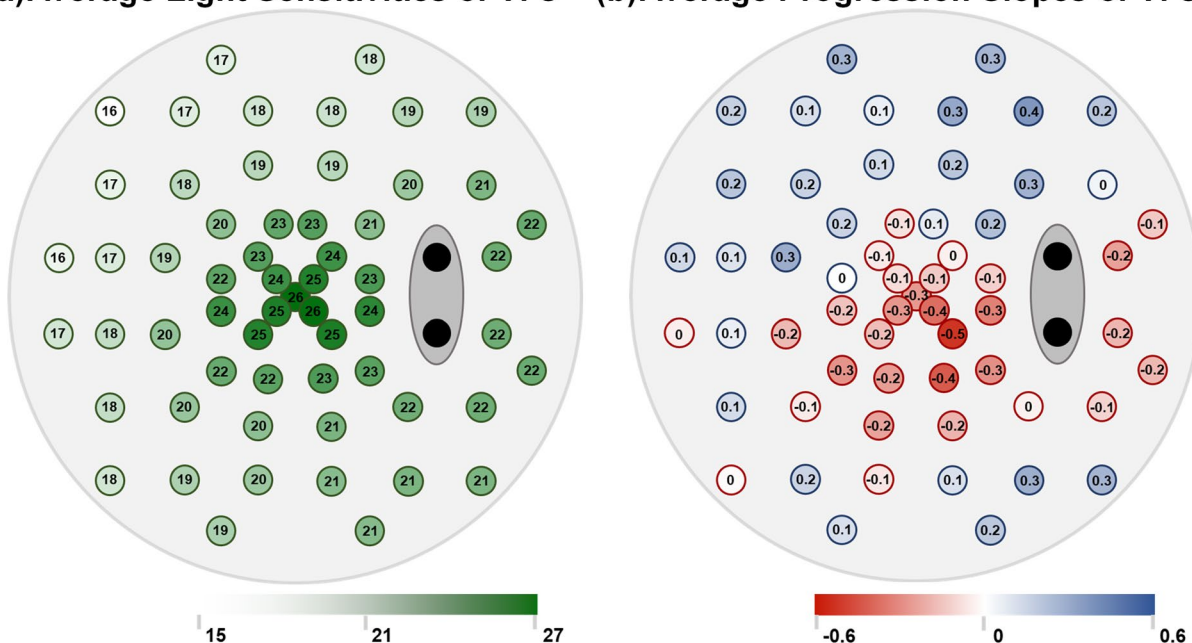
(a) Average Light Sensitivities of VFs **(b) Average Progression Slopes of VFs**

Fig. 4 Statistics characteristics of visual field (VF) values of the GRAPE dataset. (a) Average light sensitivity values of each location of all follow-up VFs. (b) Average progression slopes of each location of VFs.

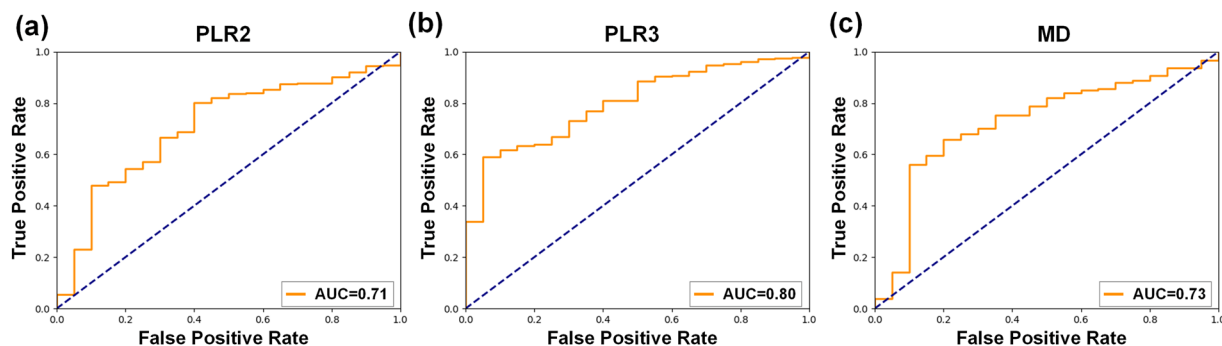


Fig. 5 The Receiver Operating Characteristic Curves (ROC) of visual field (VF) progression prediction classification model on the GRAPE dataset. (a) The ROC of VF defined by PLR2. (b) The ROC of VF defined by PLR3. (c) The ROC of VF defined by MD.

Baseline model validation for VF estimation. SAP requires the effective cooperation and reliability of the patients with a relatively lengthy test time, and the results are subjective that affected by the test-retest variability^{13–15}. Structural images or parameters have been used to estimate VF values in previous studies by DL models^{16–19}. These studies explore the relationship between structure damage and function loss as well. The CFPs and VFs in our database can be utilized in this kind of VF estimation model. A baseline model based on ResNet-50 is proposed for example as below.

In this DL model, the CFPs, ROI of CFPs, ROI of CFPs with OD/OC segmentation of each eye at baseline were considered as input, while the values of VFs at each location as output. The input images were resized to $224 \times 224 \times 3$.

We trained the same CNN, ResNet-50, with 50 epochs and a batch size of 16. The mean squared error was used as the loss function. The L2 regularization of 10^{-5} was imposed with Adam optimizer, and annealed cosine was used to update the learning rate from 10^{-4} to 10^{-5} . Gradient clipping was carried out to restrict the ℓ_2 -norm of the gradient to 1.0. The mean absolute error (MAE), root mean square error (RMSE), and R-squared (R^2), for measuring the difference between values predicted by the model and ground truth values, were used as metrics to evaluate the model estimating ability.

The MAE between predicted values and true values of each point are shown in Fig. 6, and the MAE of VF estimation from CFPs, ROI of CFPs, and ROI with OD/OC segmentation are shown in Fig. 6a–c, respectively. The prediction error is minimum at the lower part of the paracentral field, and maximum at the nasal field. The MAE, RMSE and R^2 are listed in Table 3. The MAE of the models of CFPs, ROI of CFPs, and ROI with OD/

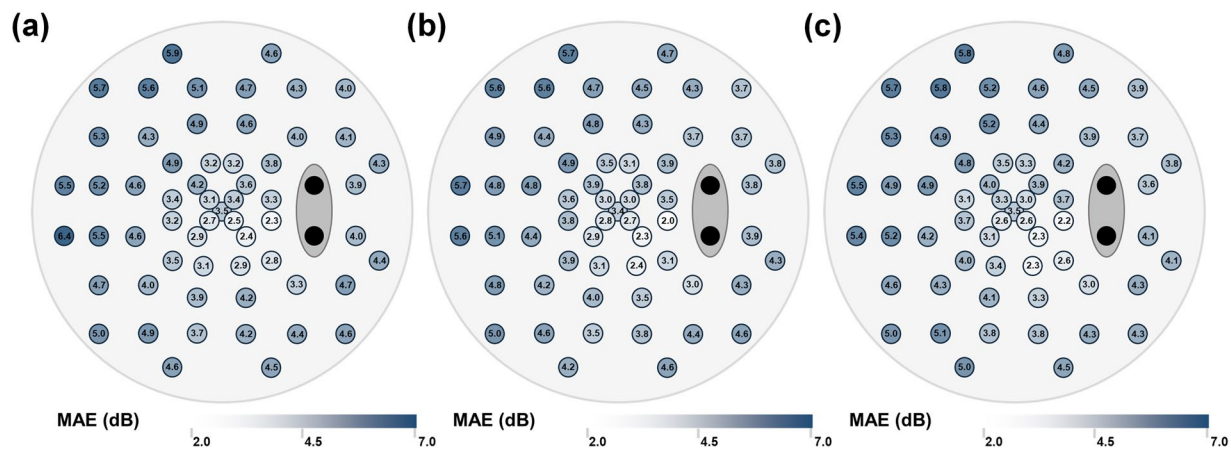


Fig. 6 The mean absolute Error (MAE) of predicted visual field (VF) at each point. (a) The MAE of VF estimated from color fundus photographs (CFPs). (b) The MAE of VF estimated from region of interest (ROI) of CFPs. (c) The MAE of VF estimated from ROI of CFPs with optic disc and optic cup segmentation.

| | RMSE | MAE | R ² |
|-----------------------------|-------|-------|----------------|
| CFPs | 5.627 | 4.143 | 0.263 |
| ROI of CFPs | 5.475 | 4.029 | 0.306 |
| ROI with OD/OC segmentation | 5.596 | 4.107 | 0.272 |

Table 3. The VF estimation error for baseline model on the GRAPE dataset.

OC segmentation is 4.143, 4.029, 4.107 respectively. It indicates that our model has the best predictive ability with ROI of CFPs, possibly because the features on ONH could be recognized by the model.

Limitations and discussion.

There are some limitations of the GRAPE dataset. First, the bias of data selection is a shortcoming of this dataset. Only patients with OAG or ACG were included (mostly OAG), and the relatively low MD (-7.21 dB) meant more mild VF defects ($MD \geq -6$ dB, stipulated by HPA criteria)³⁷ patients were selected. Although some patients with moderate (-12 dB \leq MD \leq -6 dB) and severe ($MD \leq -12$ dB) VF loss were excluded due to the VF measurement strategy and glaucoma surgery. This tendentious data inclusion would limit the generalizability of algorithm development. The performance of AI networks trained by this dataset may be affected in predicting the progression of other types of glaucoma or severe VF loss. We will consider collecting more severe VF data in the future.

Second, the inclusion data of 263 eyes with 1115 follow-up records is relatively small number for the dataset. The low number of patients was mainly owing to the setting of patient privacy, the selection of glaucoma types, the integrality of data, and the exclusion of combing with other serious eye diseases. It resulted in a small number of progressors defined by PLR3 and the possibility of overfitting of models. In the actual AI model development, the data augmentation could be a way to ameliorate the performance impact of this flaw.

Third, the detailed IOP control medications information, such as particular drug name and dosage, was not listed in this dataset. The EHR system had been updated for many times and other hospitals may dispense IOP control medicines during long the follow-up time span, leading to the information of medicines missing. To minimize the impact of treatment methods during the follow-up, we excluded the patients who had undergone glaucoma surgery to ensure that they were essentially medically-controlled.

In general, the GRAPE dataset could be used for prognostic prediction in glaucoma management and VF estimation for structure–function relationship exploration, that could advance the development of computer aided telemedicine in glaucoma.

Usage Notes

The GRAPE dataset described in this article can be downloaded through the link mentioned before. Our dataset consists of multi-modal clinical information at the baseline and longitudinal VFs, CFPs and IOP during the follow-up of glaucoma patients. The GRAPE dataset is recommended to be used to develop AI models in prognosis evaluation and VF estimation. Besides, other prediction models, such as traditional ML models, could also be applied on this dataset. Users should properly cite this article and acknowledge the contributions in their study.

Code availability

The GRAPE dataset can be downloaded at the Figshare as mentioned above³⁶. The codes of drawing annotations on ROI images are saved as the python file “draw.py”.

The DL models applied in “Technical Validation” are not the as a part of the GRAPE dataset. We uploaded these models at the Figshare as two separated parts, “Baseline model validation for VF estimation” and “Baseline model validation for VF progression prediction”, that correspond to the 2 chapters. The parameters tuning is detailed in the article.

Received: 5 February 2023; Accepted: 28 July 2023;

Published online: 05 August 2023

References

- Jonas, J. B. *et al.* Glaucoma. *The Lancet*. **390**, 2183–2193, [https://doi.org/10.1016/s0140-6736\(17\)31469-1](https://doi.org/10.1016/s0140-6736(17)31469-1) (2017).
- Weinreb, R. N. & Khaw, P. T. Primary open-angle glaucoma. *The Lancet*. **363**, 1711–1720, [https://doi.org/10.1016/s0140-6736\(04\)16257-0](https://doi.org/10.1016/s0140-6736(04)16257-0) (2004).
- Tham, Y. C. *et al.* Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. **121**, 2081–2090, <https://doi.org/10.1016/j.ophtha.2014.05.013> (2014).
- Weinreb, R. N., Aung, T. & Medeiros, F. A. The pathophysiology and treatment of glaucoma: a review. *JAMA* **311**, 1901–1911, <https://doi.org/10.1001/jama.2014.3192> (2014).
- Jammal, A. A. *et al.* Impact of Intraocular Pressure Control on Rates of Retinal Nerve Fiber Layer Loss in a Large Clinical Population. *Ophthalmology*. **128**, 48–57, <https://doi.org/10.1016/j.ophtha.2020.06.027> (2021).
- Stein, J. D., Khawaja, A. P. & Weizer, J. S. Glaucoma in Adults—Screening, Diagnosis, and Management. *JAMA*. **325**, <https://doi.org/10.1001/jama.2020.21899> (2021).
- Scuderi, G. L., Cesareo, M., Perdicchi, A. & Recupero, S. M. Standard automated perimetry and algorithms for monitoring glaucoma progression. *Progress in Brain Research*. **173**, 77–99, [https://doi.org/10.1016/S0079-6123\(08\)01107-2](https://doi.org/10.1016/S0079-6123(08)01107-2) (2008).
- Jin, K. & Ye, J. Artificial intelligence and deep learning in ophthalmology: Current status and future perspectives. *Advances in Ophthalmology Practice and Research*. **2**, <https://doi.org/10.1016/j.aopr.2022.100078> (2022).
- Daneshvar, R. *et al.* Prediction of Glaucoma Progression with Structural Parameters: Comparison of Optical Coherence Tomography and Clinical Disc Parameters. *Am J Ophthalmol*. **208**, 19–29, <https://doi.org/10.1016/j.ajo.2019.06.020> (2019).
- Dixit, A., Yohannan, J. & Boland, M. V. Assessing Glaucoma Progression Using Machine Learning Trained on Longitudinal Visual Field and Clinical Data. *Ophthalmology*. **128**, 1016–1026, <https://doi.org/10.1016/j.ophtha.2020.12.020> (2021).
- Li, F. *et al.* A deep-learning system predicts glaucoma incidence and progression using retinal photographs. *J Clin Invest*. **132**, <https://doi.org/10.1172/JCI157968> (2022).
- Shon, K., Sung, K. R. & Shin, J. W. Can Artificial Intelligence Predict Glaucomatous Visual Field Progression? A Spatial-Ordinal Convolutional Neural Network Model. *Am J Ophthalmol*. **233**, 124–134, <https://doi.org/10.1016/j.ajo.2021.06.025> (2022).
- Lee, M., Zulauf, M. & Caprioli, J. The Influence of Patient Reliability on Visual Field Outcome. *Am J Ophthalmol*. **117**, 756–761, [https://doi.org/10.1016/s0002-9394\(14\)70318-6](https://doi.org/10.1016/s0002-9394(14)70318-6) (1994).
- Gardiner, S. K., Swanson, W. H., Goren, D., Mansberger, S. L. & Demirel, S. Assessment of the Reliability of Standard Automated Perimetry in Regions of Glaucomatous Damage. *Ophthalmology*. **121**, 1359–1369, <https://doi.org/10.1016/j.ophtha.2014.01.020> (2014).
- Bengtsson, B. & Heijl, A. False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? *Invest Ophthalmol Vis Sci*. **41**, 2201–2204 (2000).
- Zhu, H. *et al.* Predicting visual function from the measurements of retinal nerve fiber layer structure. *Invest Ophthalmol Vis Sci*. **51**, 5657–5666, <https://doi.org/10.1167/iovs.10-5239> (2010).
- Guo, Z. *et al.* Optical Coherence Tomography Analysis Based Prediction of Humphrey 24-2 Visual Field Thresholds in Patients With Glaucoma. *Invest Ophthalmol Vis Sci*. **58**, 3975–3985, <https://doi.org/10.1167/iovs.17-21832> (2017).
- Christopher, M. *et al.* Deep Learning Approaches Predict Glaucomatous Visual Field Damage from OCT Optic Nerve Head En Face Images and Retinal Nerve Fiber Layer Thickness Maps. *Ophthalmology*. **127**, 346–356, <https://doi.org/10.1016/j.ophtha.2019.09.036> (2020).
- Kihara, Y. *et al.* Policy-Driven, Multimodal Deep Learning for Predicting Visual Fields from the Optic Disc and OCT Imaging. *Ophthalmology*. **129**, 781–791, <https://doi.org/10.1016/j.ophtha.2022.02.017> (2022).
- Almazroa, A. *et al.* Retinal fundus images for glaucoma analysis: the RIGA dataset. in *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*. 10579, <https://doi.org/10.1117/12.2293584> (2018).
- Zhang, Z. *et al.* ORIGA(-light): An Online Retinal Fundus Image Database for Glaucoma Analysis and Research. in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 3065–3068, <https://doi.org/10.1109/IEMBS.2010.5626137> (2010).
- Sivaswamy, J. *et al.* DRISHTI-GS: Retinal Image Dataset for Optic Nerve Head (ONH) Segmentation. in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. 53–56, <https://doi.org/10.1109/ISBI.2014.6867807> (2014).
- Orlando, J. I. *et al.* REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. *Medical Image Analysis*. **59**, <https://doi.org/10.1016/j.media.2019.101570> (2020).
- Kovalyk, O. *et al.* PAPIA: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment. *Sci Data*. **9**, 291, <https://doi.org/10.1038/s41597-022-01388-1> (2022).
- Montesano, G., Chen, A., Lu, R., Lee, C. S. & Lee, A. Y. UWHVF: A Real-World, Open Source Dataset of Perimetry Tests From the Humphrey Field Analyzer at the University of Washington. *Transl Vis Sci Technol*. **11**, 2, <https://doi.org/10.1167/tvst.11.1.1> (2022).
- Chen, J. S., Lin, W. C., Yang, S., Chiang, M. F. & Hribar, M. R. Development of an Open-Source Annotated Glaucoma Medication Dataset From Clinical Notes in the Electronic Health Record. *Transl Vis Sci Technol*. **11**, 20, <https://doi.org/10.1167/tvst.11.11.20> (2022).
- Gordon, M. O. *et al.* Evaluation of a Primary Open-Angle Glaucoma Prediction Model Using Long-term Intraocular Pressure Variability Data: A Secondary Analysis of 2 Randomized Clinical Trials. *JAMA Ophthalmol*. **138**, 780–788, <https://doi.org/10.1001/jamaophthalmol.2020.1902> (2020).
- Prum, B. E. Jr. *et al.* Primary Open-Angle Glaucoma Preferred Practice Pattern[®] Guidelines. *Ophthalmology*. **123**, 41–111, <https://doi.org/10.1016/j.ophtha.2015.10.053> (2016).
- Prum, B. E. Jr. *et al.* Primary Angle Closure Preferred Practice Pattern[®] Guidelines. *Ophthalmology*. **123**, 1–40, <https://doi.org/10.1016/j.ophtha.2015.10.049> (2016).
- Advanced Glaucoma Intervention Study. *Ophthalmology*. **101**, 1445–1455, [https://doi.org/10.1016/s0161-6420\(94\)31171-7](https://doi.org/10.1016/s0161-6420(94)31171-7) (1994).
- Newkirk, M. R., Gardiner, S. K., Demirel, S. & Johnson, C. A. Assessment of false positives with the Humphrey Field Analyzer II perimeter with the SITA Algorithm. *Invest Ophthalmol Vis Sci*. **47**, 4632–4637, <https://doi.org/10.1167/iovs.05-1598> (2006).
- Sample, P. A. *et al.* Imaging and Perimetry Society standards and guidelines. *Optometry and vision science: official publication of the American Academy of Optometry*. **88**, 4–7, <https://doi.org/10.1097/OPX.0b013e3181fc3735> (2011).
- Vesti, E., Johnson, C. A. & Chauhan, B. C. Comparison of different methods for detecting glaucomatous visual field progression. *Invest Ophthalmol Vis Sci*. **44**, 3873–3879, <https://doi.org/10.1167/iovs.02-1171> (2003).
- Saeedi, O. J. *et al.* Agreement and Predictors of Discordance of 6 Visual Field Progression Algorithms. *Ophthalmology*. **126**, 822–828, <https://doi.org/10.1016/j.ophtha.2019.01.029> (2019).

35. Marin-Franch, I. & Swanson, W. H. The visualFields package: a tool for analysis and visualization of visual fields. *J Vis.* **13**, <https://doi.org/10.1167/13.4.10> (2013).
36. Huang, X. *et al.* GRAPE: A multi-modal dataset of longitudinal follow-up visual field and fundus images for glaucoma management, *Figshare*, <https://doi.org/10.6084/m9.figshare.c.6406319.v1> (2023).
37. Hodapp, E., Parrish, R. K. & Anderson, D. R. *Clinical Decisions in Glaucoma*. St. Louis, MO: Mosby (1993).

Acknowledgements

This work has been financially supported by National Natural Science Foundation of China (grant number U20A20386), National Key Research and Development Program of China (grant number 2019YFC0118400), Natural Science Foundation of China (grant number 82201195), Key Research and Development Program of Zhejiang Province (grant number 2019C03020), and Clinical Medical Research Center for Eye Diseases of Zhejiang Province (grant number 2021E50007).

Author contributions

X.H.: methodology, data acquisition and analysis, manual segmentation and writing. X.K.: data analysis, image process and algorithm design. Z.S.: data analysis, image process and algorithm design. J.O.: image process and algorithm design. Y.L.: algorithm design. K.J.: conceptualization, project administration and writing. J.Y.: conceptualization, funding acquisition and reviewing. All authors have read and approved the final manuscript. J.Y. is guarantor.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.J. or J.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023