Original Article

A multi-center study of prediction of macular hole status after vitrectomy and internal limiting membrane peeling by a deep learning model

Yijun Hu[1,2,3#], Yu Xiao[1#], Wuxiu Quan[4#], Bin Zhang[4], Yuqing Wu[5], Qiaowei Wu[1], Baoyi Liu[1], Xiaomin Zeng[1], Ying Fang[1], Yu Hu[6], Songfu Feng[7], Ling Yuan[6], Tao Li[5], Hongmin Cai[4], Honghua Yu[1]

[1]Guangdong Eye Institute, Department of Ophthalmology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, The Second School of Clinical Medicine, Southern Medical University, Guangzhou, China; [2]Aier Institute of Refractive Surgery, Refractive Surgery Center, Guangzhou Aier Eye Hospital, Guangzhou, China; [3]Aier School of Ophthalmology, Central South University, Changsha, China; [4]School of Computer Science and Engineering, South China University of Technology, Guangzhou, China; [5]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China; [6]Department of Ophthalmology, The First Affiliated Hospital of Kunming Medical University, Kunming, China; [7]Department of Ophthalmology, Zhujiang Hospital of Southern Medical University, Guangzhou, China

Contributions: (I) Conception and design: Y Hu, H Yu, H Cai, T Li, L Yuan; (II) Administrative support: Y Hu, H Yu, H Cai; (III) Provision of study materials or patients: H Yu, H Cai, T Li, L Yuan, S Feng; (IV) Collection and assembly of data: Y Xiao, Y Fang, Y Wu, Y Hu, S Feng, X Zeng; (V) Data analysis and interpretation: Y Xiao, W Quan, B Zhang, Y Hu, Q Wu, B Liu, H Cai, H Yu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work as co-first authors.

Correspondence to: Honghua Yu. Guangdong Eye Institute, Department of Ophthalmology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, The Second School of Clinical Medicine, Southern Medical University, No. 106 Zhongshan Er Road, Yuexiu District, Guangzhou 510080, China. Email: yuhonghua@gdph.org.cn; Hongmin Cai. School of Computer Science and Engineering, South China University of Technology, Higher Education Mega Center, Panyu District, Guangzhou 510006, China. Email: hmcai@scut.edu.cn. Tao Li, State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, 7 Jinsui Road, Guangzhou 510000, China. Email: hxydlitao@163.com.

Background: To develop a deep learning (DL) model for prediction of idiopathic macular hole (MH) status after vitrectomy and internal limiting membrane peeling (VILMP) based on optical coherence tomography (OCT) images from four ophthalmic centers.

Methods: Eyes followed up at 1 month after VILMP for full-thickness MH were included. In the internal training set, 920 preoperative macular OCT images (as the input) and post-operative status of MH (closed or open, as the output) of 256 eyes from two ophthalmic centers were used to train the DL model using VGG16 algorithm. In the external validation set, 72 preoperative macular OCT images of 36 MH eyes treated by VILMP from another two ophthalmic centers were used to validate the prediction accuracy of the DL model.

Results: In internal training, the mean of overall accuracy for prediction of MH status after VILMP was 84.6% with a mean area under the receiver operating characteristic (ROC) curve (AUC) of 91.04% (sensitivity 85.37% and specificity 81.99%). In external validation, the overall accuracy of predicting MH status after VILMP was 84.7% with an AUC of 89.32% (sensitivity 83.33% and specificity 87.50%). The heatmaps showed that the area critical for prediction was at the central macula, mainly at the MH and its adjacent retina.

Conclusions: The DL model trained by preoperative macular OCT images can be used to predict postoperative MH status after VILMP. The prediction accuracy of our DL model has been validated by multiple ophthalmic centers.

Keywords: Deep learning (DL); macular hole; optical coherence tomography; clinical prediction model

Submitted Feb 21, 2020. Accepted for publication Sep 17, 2020.
doi: 10.21037/atm-20-1789
View this article at: http://dx.doi.org/10.21037/atm-20-1789

Introduction

Idiopathic macular hole (MH) is a discontinuation of the neurosensory retina at the center of the macula (1,2). The mechanism of MH development is believed to be caused by pathological vitreoretinal traction at the macula (3,4). Patients with MH typically experience progressive visual impairment and metamorphopsia (5). The prevalence of MH ranges from 0.1% to 0.8% in adults over 44 years old (6-9), while the age- and sex-adjusted incidence has been reported to be 7.8 persons or 8.69 eyes per 100,000 population (10). According to the Beijing Eye Study, the prevalence of MH is 0.09%±3.04% in China (7). Older age and female gender are possible risk factors for MH development (5,10,11). Patients with full-thickness MH in one eye are also at risk of MH development in the other eye (5,12). Vitrectomy and internal limiting membrane peeling (VILMP) has been commonly used to treat MH, achieving a success rate of within 80–95% (13-17). A standard VILMP surgery includes a three-port vitrectomy (23-gauge or 25-gauge), internal limiting membrane (ILM) peeling with or without staining, and air tamponade.

Despite high success rate of MH surgery, in some cases the MH remains open after routine VILMP (18). In patients with an open MH after initial surgery, a second surgery is often required. However, the second surgery is mainly associated with high medical costs and less favorable visual outcomes compared with primary closure (19). Therefore, it is of clinical importance to identify MH in the risk of surgical failure after standard VILMP. Risk factors of unclosed MH include older age, larger base diameter, longer minimum linear dimension, longer hole duration, etc. (14,20-22). Among these factors, parameters related to MH dimensions are of significant importance (14,23).

With the development of optical coherence tomography (OCT) technology, it has been widely used for diagnosis and follow-up of MH. To better predict MH status after surgery, several OCT parameters of MH have been proposed (14,23-25). For instance, some researchers have demonstrated that preoperative MH diameter determined by OCT is related to post-operative closure of the MH (25). In another study, it has been shown that the hole form factor (HFF) of MH is associated with anatomical success rate following the initial surgical procedure (23). However, the measurements of these parameters vary between technicians and the prediction accuracy of them is not satisfactory (14).

Deep learning (DL) is state-of-the-art technology in artificial intelligence, which has sparked tremendous global interest over the last few years (26). DL systems have expert-level performance in detecting various ocular diseases including diabetic retinopathy (DR) and age-related macular degeneration (AMD) using clinical ocular images (27,28). On the other hand, DL systems also have been used for predicting treatment outcomes of ocular diseases based on pretreatment clinical images. Recently, a DL-based model developed by Gupta *et al.* has been applied to predict and monitor retinopathy of prematurity (ROP) regression after treatment (29). However, to the best of our knowledge, no DL model has been developed to predict MH status based on the preoperative macular OCT images. If available, such a DL model could help vitreoretinal surgeons to accurately select patients, who are the most likely to have unclosed MH after routine VILMP, and a new surgical technique such as inverted ILM flap can be suggested to those patients (30).

This study aimed to develop a DL model to automatically predict MH status after routine VILMP based on preoperative macular OCT images from a multi-center population.

Methods

Patients and preparation of datasets

Eyes followed up at 1 month after VILMP for MH were retrospectively included in this study. Eyes with a macular hole caused by known etiologies such as trauma, macular edema, epiretinal membrane, high myopia, retinal detachment or retinoschisis were excluded. All the eyes received comprehensive ophthalmologic examinations including best-corrected visual acuity (BCVA), slit-lamp biomicroscope anterior segment and fundus examination, intraocular pressure (IOP) measurement, and SD-OCT scanning (Spectralis; Heidelberg Engineering, Heidelberg, Germany) before and after VILMP. The study was conducted according to the Declaration of Helsinki (as revised in 2013) and was approved by the Institutional Review Board of GPPH (No. GDREC2020067H). Informed consent was taken from all the patients. A total of 920 preoperative macular OCT images of 256 eyes from the Department of Ophthalmology, Guangdong Provincial People's Hospital (GDPH, 213 images of 54 eyes) and the Zhongshan Ophthalmic Center (707 images of 202 eyes) were collected for training and internal validation of the proposed DL model. A total of 72 preoperative macular OCT images of 36 eyes from the Department of Ophthalmology, Zhujiang Hospital of Southern Medical University (ZHSMU, 44 images of 27 eyes) and the Department of Ophthalmology, the First Affiliated Hospital of Kunming Medical University (FAHKMU, 28 images of 9 eyes) were included in the external validation dataset.

OCT examination and IMH status labeling

OCT examinations were performed by experienced technicians. A custom of 20°×20° volume acquisition protocol was used to obtain a set of high-speed scans from each eye. With this protocol, 25 horizontal and central vertical cross-sectional B-scan images were obtained, which consisted of 512 A-scans each. The image through the fovea was determined by simultaneous evaluation of the red-free image on the computer monitor of the OCT scanner. To establish a standardized image format of the dataset for subsequent training and validation, all scans were saved in the TIFF format. Before training, all OCT images had undergone a layered labeling system consisting of multiple layers of trained graders with increasing expertise for verification and correction of image labels. The first layer of graders conducted quality control and excluded images with low quality. The images taken from improper positioning, with low signals, or with strong motion artifacts causing misalignment and blurring of sections were excluded. The second layer of graders consisted of two Chinese board-certified ophthalmologists (YH, YX) who independently labeled the MH status (closed or open) according to the OCT images that had passed the first layer. MH closure was defined as restoration of continuity of neurosensory retina at the central fovea in all the post-operative OCT scans. Finally, the third layer grader who was a senior retinal specialist (YH) verified the true labels of MH status and generated the final decision if there was any discordance between the 2 ophthalmologists. The MH status verified by the third layer grader was used as the true label of each OCT image.

Training and validation of DL-based model

The raw OCT images were first preprocessed to normalize the input data. The saturated pixels with an intensity value of 255 were removed from the raw OCT images and the retinal layers based on smooth pixel intensity were cropped. The OCT images were then resized into 224*224 pixels based on the requirements of Visual Geometry Group (VGG) 16 (Department of Engineering Science, the University of Oxford, Oxford, UK) (31). The VGG16 with 16 convolutional layers was used as the benchmark DL-based model in the whole experiments, wherein the OCT images (B-scans containing the macular center) were the input, and the status of MH after VILMP were the output (*Figure 1*).

The training started with multiple iterations with a batch size of 16 images, with the initial learning rate of 0.001 and stopped at 15 epochs. For each training iteration, a cross-entropy loss function was used as the objective loss function to update the optimization parameters, and a stochastic gradient descent (SGD) algorithm with Nesterov momentum term was used to optimize a pre-defined loss of function to train neuron weights via back propagation. At every epoch, the performance of the convolutional neural network (CNN) was assessed using the validation dataset. The VGG16 model was a binary classification model output by the Softmax classifier, and 2 nodes were used in the last layer to generate predictions.

In the training dataset, there were 96.88% of patients who had more than one OCT image per eye, with an average of 3.6 OCT images per eye. To determine the most appropriate hyperparameters for our final model, the popular ten-fold cross-validation scheme was used on the training dataset, which was divided into ten portions randomly. To obtain fully independent folds, python code was used to read the labels and the corresponding index positions of all images. The code was then used to remove the repeated names of different index position images of the same patient (in other words, only one patient name was kept in several images of the same patient). Therefore, the partition was carried out at the patient-level, and the original index position images belonging to the same

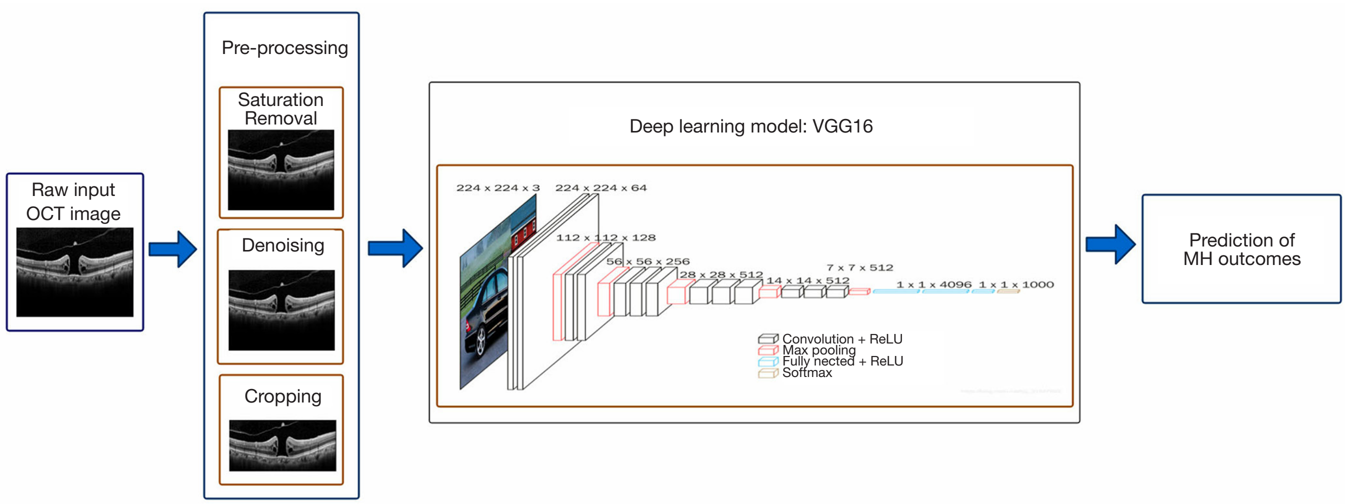

Figure 1 Demonstration of construction of the deep learning model. A deep learning model was trained to predict MH status (closed or open) after surgery using VGG16 network. MH, macular hole; OCT, optical coherence tomography; VGG16, Visual Geometry Group 16 Layers.

patient were divided into the same partition to ensure the completely independent folds. In each run, nine portions of the dataset were employed to train the DL model, and the other one was used for model testing to facilitate parameter selection and tuning. The experiments were conducted until each portion was tested. Using the above-chosen hyperparameters, the DL system was re-trained using the entire training dataset, and its predictive performance was evaluated based on our independent external validation dataset, in which the data were not involved in the development of DL-based method.

Statistical analysis

After training, validating and testing our DL model, we could obtain the predictive output of each OCT image. The exact breakdown of performance regarding the correlation of predicted labels obtained from our DL-based model with true labels was depicted as confusion matrices. The confusion matrices were used to calculate the overall accuracy of prediction of MH status after VILMP.

The area under the receiver operating characteristic (ROC) curve (AUC) was used to evaluate the accuracy of DL model in predicting the MH status after surgery. A series of true positive rates (TPRs) and false positive rates (FPRs) were obtained to form ROC curves. The TPR was also known as sensitivity, and the FPR results were obtained by subtracting the specificity value from 1. The optimal cut-off point was obtained by using the highest Youden's index (sensitivity+specificity−1), and the corresponding optimal sensitivity and specificity values were recorded. The overall accuracy and the AUC in internal validation were presented as mean and 95% confidence interval (CI).

Visualization method for the proposed DL-based model

To visualize the critical area in OCT images that was highly correlated with MH status after VILMP, Gradient-weighted Class Activation Mapping (Grad-CAM), an approach proposed by Ramprasaath R. Selvaraju *et al.*, was used to aid interpretation of the results and increase the model transparency (32). The gradient indicated the contribution of each spot of the feature map to the output probability value. The region with the largest gradient had the greatest impact on the output probability value. The highlighted region in the heatmap represented the part of the OCT image most critical for accurately predicting the MH status after VILMP.

Results

Demographics of the eyes included were shown in *Table 1*. There were 208 eyes with closed MH and 84 eyes with open MH at 1-month visit (to balance the number of closed and open MH, more eyes with an open MH were deliberately included). In internal validation, the mean overall accuracy

Table 1 Patient demographics

Characteristics	All eyes	Internal validation	External validation
Number of eyes	292	256	36
Age, mean (SD), years	60.36 (10.83)	60.35 (10.31)	60.44 (14.58)
Sex, females, n (%)	185 (63.37)	165 (64.45)	20 (55.56)
Duration of symptoms, mean (SD), months	7.20 (12.35)	7.11 (11.98)	7.89 (15.25)
Preoperative BCVA, mean (SD), logMAR	1.04 (0.43)	1.04 (0.44)	1.01 (0.32)
Number of images	992	920	72
Images with a closed MH, n (%)	681 (68.65)	633 (68.80)	48 (66.67)

MH, macular hole; SD, standard deviation; BCVA, best-corrected visual acuity; logMAR, the logarithm of minimal angle resolution.

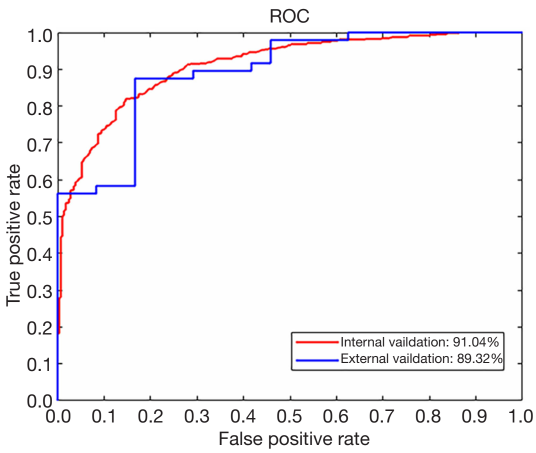

Figure 2 The receiver operating characteristic (ROC) curves of the deep learning model. The ROC curves for the prediction task in the internal validation set (red line) and external validation set (blue line). The area under ROC curves (AUC) for the internal validation set and the external validation set was 91.04% and 89.32%, respectively.

for predicting the MH status after VILMP was 84.6% (95% CI: 82.89–86.31%), with a mean area under the receiver operating characteristic (ROC) curve (AUC) of 91.04% (95% CI: 90.09–91.99%, sensitivity 85.37% and specificity 81.99%). In external validation, the overall accuracy for predicting the MH status after VILMP was 84.7%, with an AUC of 89.32% (sensitivity 83.33% and specificity 87.50%). The prediction accuracy of MH closure and opening was 85.4% and 83.3%, respectively (*Figures 2,3*). Heatmaps showed that the area critical for prediction was at the central macula, mainly at the MH and its adjacent retina, suggesting that the DL-based model could assist in successfully identifying the pathologic region that was the most critical for accurate prediction of MH status after VILMP (*Figure 4*).

Discussion

To the best of our knowledge, the present study is the first to predict the anatomical outcome of standard MH surgery based on preoperative macular OCT images from a multicenter dataset. The results show that the accuracy of our DL model in predicting MH status (i.e., closed or open) is promising. The performance of DL model is as follows: a mean overall accuracy of 84.6% and a mean AUC of 91.04% in internal validation, and an overall accuracy of 84.7% and an AUC of 89.32% in external validation. Moreover, heatmaps of the DL model show that the area highlighted is the most predominant pathologic region of the MH. These findings indicate that our DL model is capable of successfully predicting the MH status by accurately recognizing the critical pathologic region in preoperative OCT images.

The mechanisms of MH development mainly involve anterior and tangential vitreoretinal traction to the fovea (3). Accordingly, release of vitreoretinal traction by VILMP has become a standard treatment for FTHM with success rates of VILMP vary from 80% to 95% (13-17). Despite high success rates of VILMP, the MH remains to be open after uneventful surgery in a small number of cases (18,30). Preoperative OCT parameters of MH are important predictive factors associated with MH status after VILMP surgery (14,20). Therefore, some researchers have suggested that adjusting surgical planning according to OCT parameters of the MH to increase success rates (33). These previous findings constitute the basis of

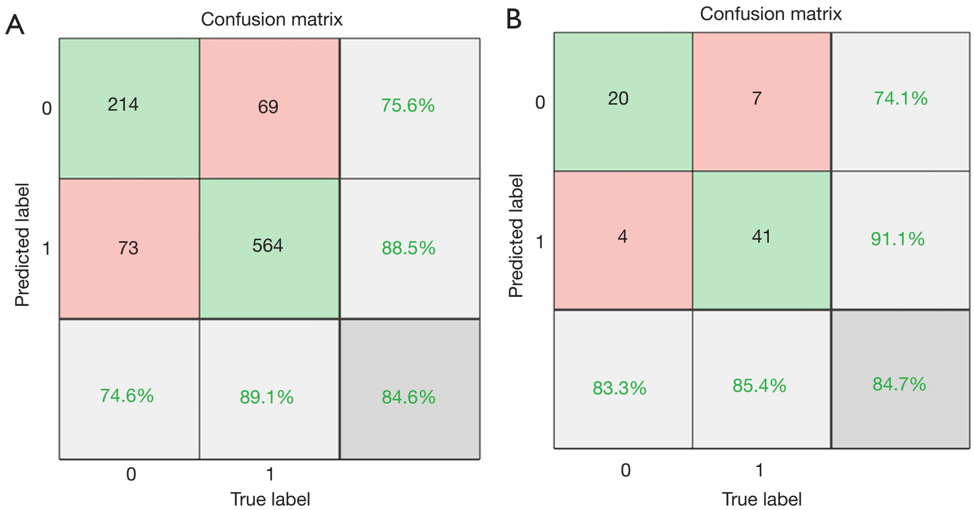

Figure 3 Confusion matrix for binary classifications using the deep learning model. Ground true labels are on the vertical axis and predicted labels are on the horizontal axis. (A) Confusion matrix for the internal validation set. The overall accuracy was 84.6%, while the accuracy for predicting macular hole closure or opening was 89.1% and 74.6%, respectively. (B) Confusion matrix for the external validation set. The overall accuracy was 84.7%, while the accuracy for predicting macular hole closure or opening was 85.4% and 83.3%, respectively.

the possibility and necessity for the prediction of MH status after VILMP surgery. However, current measurements of the MH OCT parameters used for predicting are manual, expertise-requiring and time-consuming. The measurements also vary between different technicians. It is particularly beneficial and meaningful to develop an automated system to select MH at a high risk of surgery failure after standard VILMP.

Using the DL technology, information in the preoperative OCT images associated with postoperative MH status can be automatically extracted and processed. The information is unique for each OCT image, and contains parameters that have been reported before and parameters that are unknown so far. For each OCT image, the information is unique and is integrated by DL model. By corresponding the preoperative OCT information to the postoperative MH status after repeated training, the ML model can predict the MH status based on the OCT images. The DL method we have adopted in this study (VGGNet) consists of 16 layers, mainly composing of small 3×3 convolution operations and 2×2 pooling operations (31). Therefore, stacking multiple small convolution kernels can increase the depth of the network to acquire richer image characteristics, while limiting the number of parameters. For example, by stacking three 3×3 convolutional layers instead of using a single 7×7 convolutional layer (i.e., AlexNet), some limitations can be overcome. Firstly, it combines three nonlinear functions rather than one, making the decision function more discriminative and representative. Secondly, the number of parameters is reduced, while the receptive field remains unchanged. In addition, the use of small convolution kernels also plays a role in regulating and improving the effectiveness of different convolution kernels.

DL studies based on OCT images mostly concentrate on image segmentation, involving complex feature selection and extraction. In addition, introducing a minor error in the segmentation can lead to classification error (34). However, this problem can be avoided in the present study by training the DL model to learn the prediction features directly from the OCT images. Therefore, compared to CNN with shallower architectures, our DL-based model can automatically learn richer and more discriminative OCT image features to present an accurate prediction.

Our deep neural network also contains the batch normalization (BN) layer and the dropout layer. The BN layer can pull the distribution of neuron input value of each neural network layer back to the standard normal distribution with a mean of 0 and a variance of 1. Therefore, the BN layer not only improves the training speed, but also prevents gradient exploding and gradient vanishing. Moreover, the BN layer can avoid overfitting to a certain extent. In addition, the dropout layer, placed after the fully connected layer, is used as a trick for training deep neural

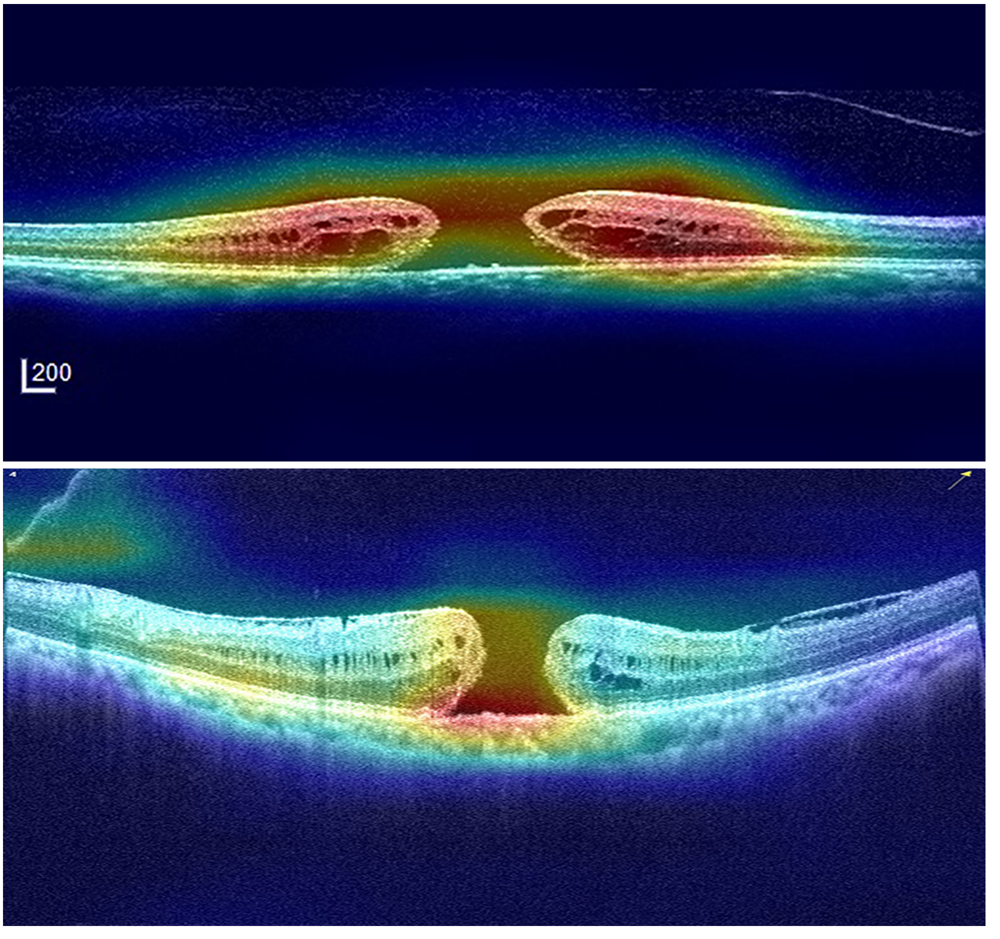

Figure 4 Heatmaps highlighting the pathological area highly correlated with macular hole status after surgery. The heatmaps were generated by Gradient-weighted Class Activation Mapping (Grad-CAM). The heatmaps demonstrate the critical area in optical coherence tomography images that were highly correlated with an accurate prediction of macular hole status after surgery.

networks. In each training batch, randomly dropping out half of the feature detectors (assigning half of the hidden layer nodes to 0) can reduce the interaction between the feature detectors (hidden layer nodes), and reduces the complex co-adaptation relationship between the neurons. This approach forces the network more robustly and significantly avoided overfitting.

A major strength of our study is the heterogeneity of the datasets. The DL-based model is trained by OCT images from two different ophthalmic centers, and it has been validated by an independent dataset from another two ophthalmic centers. The heterogeneity of the datasets is very important for the development and validation of DL model (35). Data from patients with different backgrounds and demographics can help train the DL-based model to learn the "true features" critical for making an accurate prediction. Successful validation of the DL models by different external datasets would be more convincing for predicting the ability and popularity of the models. While we increase the heterogeneity of our datasets, we have only included patients with idiopathic MH but not MH secondary to known etiologies such as trauma. This is because MH caused by known etiologies might have different pathogenesis and prognostic factors from idiopathic MH. The inclusion of these cases may reduce the efficacy of training, leading to poor performance of the DL-based model. Thus, to ensure the prediction accuracy of the DL-based model, MH secondary to known etiologies is excluded for the current study.

Another artificial intelligence technology called machine learning (ML) also has been proposed to predict treatment outcomes of some ocular diseases, such as AMD and macular edema. Unlike DL models, ML models use text information instead of images as input. In a recent study, information extracted from OCT images taken during the upload phase of 3 anti-VEGF injections was incorporated into an ML-based model to predict recurrence of macular edema associated with retinal vein occlusion. The results were encouraging with an AUC of 0.76–0.83 (36). However, unlike the severity of macular edema, the parameters of MH morphology are often manually measured, restricting the use of ML model in predicting the MH status after surgery. In the present study, macular OCT image is used as a whole to train the DL-based model. The DL-based model can automatically process and analyze anatomical parameters of the MH. Hence, the problems associated with manual measurement are avoided.

However, there are several limitations in this study. One limitation is the small sample size. One may argue that an external validation with 36 eyes is not convincing enough. Although the sample is not large considering this is a DL-based study, we have shown excellent accuracy of our DL-based model in MH status prediction after VILMP. Moreover, the eyes in the external validation set are obtained from two ophthalmic centers different from the internal validation set, suggesting satisfactory adaptability of the DL model. On the other hand, this is a preliminary study conducted to evaluate the possibility of predicting postoperative MH status using DL model. Further studies with a larger sample size from multiple centers are warranted to validate the results of the present study.

In conclusion, the proposed DL-based model has demonstrated high accuracy and transparency in automated prediction of MH status after VILMP based on preoperative macular OCT images. Our results indicate the feasibility of automatic prediction of MH status after routine vitreoretinal surgery. In case of further improving and fine-tuning, the proposed DL-based model can be used to select patients with a MH that is very unlikely to be closed using routine vitreoretinal surgery and more progressive surgical methods can be suggested to these patients.

Acknowledgments

The authors thank the School of Computer Science and Engineering, South China University of Technology for their technical assistance with the deep learning system; the Zhongshan Ophthalmic Center, the Department of Ophthalmology in Zhujiang Hospital of Southern Medical University and the Department of Ophthalmology in the First Affiliated Hospital of Kunming Medical University for contributing OCT images for internal and external validation.

Funding: This study was supported by the National Natural Science Foundation of China (No. 81870663 and No. 61771007), the Science and Technology Program of Guangzhou (No. 202002030074), the Outstanding Young Talent Trainee Program of Guangdong Provincial People's Hospital (No. KJ012019087), the GDPH Scientific Research Funds for Leading Medical Talents and Distinguished Young Scholars in Guangdong Province (No. KJ012019457), the talent introduction fund of Guangdong Provincial People's Hospital (No. Y012018145), the Technology Innovation Guidance Program of Hunan Province (No. 2018SK50106), the Science and Technology Planning Project of Guangdong Province (No. 2017B020226004), the Health and Medical Collaborative Innovation Project of Guangzhou City (No. 201803010021), and the Fundamental Research Fund for the Central Universities (No. 2017ZD051), the Science Research Foundation of Aier Eye Hospital Group (No. AM1909D2).

Footnote

Data Sharing Statement: Available at http://dx.doi.org/10.21037/atm-20-1789

Peer Review File: Available at http://dx.doi.org/10.21037/atm-20-1789

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-1789). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted according to the Declaration of Helsinki (as revised in 2013) and was approved by the Institutional Review Board of GPPH (No. GDREC2020067H). Informed consent was taken from all the patients.

References

1. Barak Y, Ihnen MA, Schaal S. Spectral domain optical coherence tomography in the diagnosis and management of vitreoretinal interface pathologies. J Ophthalmol 2012;2012:876472.
2. Takahashi A, Yoshida A, Nagaoka T, et al. Idiopathic full-thickness macular holes and the vitreomacular interface: a high-resolution spectral-domain optical coherence tomography study. Am J Ophthalmol 2012;154:881-92.e2.
3. Gass JD. Idiopathic senile macular hole. Its early stages and pathogenesis. Arch Ophthalmol 1988;106:629-39.
4. Gass JD. Reappraisal of biomicroscopic classification of stages of development of a macular hole. Am J Ophthalmol 1995;119:752-9.
5. Chew EY, Sperduto RD, Hiller R, et al. Clinical course of macular holes: the Eye Disease Case-Control Study. Arch Ophthalmol 1999;117:242-6.
6. la Cour M, Friis J. Macular holes: classification, epidemiology, natural history and treatment. Acta Ophthalmol Scand 2002;80:579-87.
7. Wang S, Xu L, Jonas JB. Prevalence of full-thickness macular holes in urban and rural adult Chinese: the Beijing Eye Study. Am J Ophthalmol 2006;141:589-91.
8. Sen P, Bhargava A, Vijaya L, et al. Prevalence of idiopathic macular hole in adult rural and urban south Indian population. Clin Exp Ophthalmol 2008;36:257-60.
9. Thapa SS, Thapa R, Paudyal I, et al. Prevalence and pattern of vitreo-retinal diseases in Nepal: the Bhaktapur glaucoma study. BMC Ophthalmol 2013;13:9.
10. McCannel CA, Ensminger JL, Diehl NN, et al. Population-based incidence of macular holes. Ophthalmology 2009;116:1366-9.
11. Johnson MW, Van Newkirk MR, Meyer KA. Perifoveal vitreous detachment is the primary pathogenic event in idiopathic macular hole formation. Arch Ophthalmol 2001;119:215-22.
12. Kumagai K, Ogino N, Hangai M, et al. Percentage of fellow eyes that develop full-thickness macular hole in patients with unilateral macular hole. Arch Ophthalmol 2012;130:393-4.
13. Passemard M, Yakoubi Y, Muselier A, et al. Long-term outcome of idiopathic macular hole surgery. Am J Ophthalmol 2010;149:120-6.
14. Wakely L, Rahman R, Stephenson J. A comparison of several methods of macular hole measurement using optical coherence tomography, and their value in predicting anatomical and visual outcomes. Br J Ophthalmol 2012;96:1003-7.
15. Krishnan R, Tossounis C, Fung Yang Y. 20-gauge and 23-gauge phacovitrectomy for idiopathic macular holes: comparison of complications and long-term outcomes. Eye (Lond) 2013;27:72-7.
16. Elhusseiny AM, Schwartz SG, Flynn HW Jr, Smiddy WE. Long-Term Outcomes after Macular Hole Surgery. Ophthalmol Retina 2020;4:369-76.
17. Christmas NJ, Smiddy WE, Flynn Jr HW. Reopening of macular holes after initially successful repair. Ophthalmology 1998;105:1835-8.
18. Abbey AM, Van Laere L, Shah AR, Hassan TS. Recurrent macular holes in the era of small-gauge vitrectomy: A Review of Incidence, Risk Factors, and Outcomes. Retina 2017;37:921-4.
19. Valldeperas X, Wong D. Is it worth reoperating on macular holes? Ophthalmology 2008;115:158-63.
20. Salter AB, Folgar FA, Weissbrot J, et al. Macular hole surgery prognostic success rates based on macular hole size. Ophthalmic Surg Lasers Imaging 2012;43:184-9.
21. Kim Y, Kim ES, Yu SY, Kwak HW. Age-related clinical outcome after macular hole surgery. Retina 2017;37:80-7.
22. Essex RW, Kingston ZS, Moreno-Betancur M, et al. The Effect of Postoperative Face-Down Positioning and of Long- versus Short-Acting Gas in Macular Hole Surgery: Results of a Registry-Based Study. Ophthalmology 2016;123:1129-36.
23. Ullrich S, Haritoglou C, Gass C, et al. A. Macular hole size as a prognostic factor in macular hole surgery. Br J Ophthalmol 2002;86:390-3.
24. Ruiz-Moreno JM, Staicu C, Pinero DP, et al. Optical coherence tomography predictive factors for macular hole surgery outcome. Br J Ophthalmol 2008;92:640-4.
25. Ip MS, Baker BJ, Duker JS, et al. Anatomical outcomes of surgery for idiopathic macular hole as determined by optical coherence tomography. Arch Ophthalmol 2002;120:29-35.
26. Kermany DS, Goldbaum M, Cai W, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 2018;172:1122-31.e9.
27. Lu W, Tong Y, Yu Y, et al. Deep Learning-Based

Automated Classification of Multi-Categorical Abnormalities From Optical Coherence Tomography Images. Transl Vis Sci Technol 2018;7:41.
28. Li F, Chen H, Liu Z, et al. Deep learning-based automated detection of retinal diseases using optical coherence tomography images. Biomed Opt Express 2019;10:6204-26.
29. Gupta K, Campbell JP, Taylor S, et al. A Quantitative Severity Scale for Retinopathy of Prematurity Using Deep Learning to Monitor Disease Regression After Treatment. JAMA Ophthalmol 2019;137:1029-36.
30. Michalewska Z, Michalewski J, Adelman RA, et al. Inverted internal limiting membrane flap technique for large macular holes. Ophthalmology 2010;117:2018-25.
31. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556 2014.
32. Selvaraju RR, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision, 2017.
33. Yao Y, Qu J, Dong C, et al. The impact of extent of internal limiting membrane peeling on anatomical outcomes of macular hole surgery: results of a 54-week randomized clinical trial. Acta Ophthalmol 2019;97:303-12.
34. Xu X, Lee K, Zhang L, et al. Stratified Sampling Voxel Classification for Segmentation of Intraretinal and Subretinal Fluid in Longitudinal Clinical OCT Data. IEEE Trans Med Imaging 2015;34:1616-23.
35. Ting DSW, Lee AY, Wong TY. An Ophthalmologist's Guide to Deciphering Studies in Artificial Intelligence. Ophthalmology 2019;126:1475-9.
36. Vogl WD, Waldstein SM, Gerendas BS, et al. Predicting Macular Edema Recurrence from Spatio-Temporal Signatures in Optical Coherence Tomography Images. IEEE Trans Med Imaging 2017;36:1773-83.

Cite this article as: Hu Y, Xiao Y, Quan W, Zhang B, Wu Y, Wu Q, Liu B, Zeng X, Fang Y, Hu Y, Feng S, Yuan L, Li T, Cai H, Yu H. A multi-center study of prediction of macular hole status after vitrectomy and internal limiting membrane peeling by a deep learning model. Ann Transl Med 2021;9(1):51. doi: 10.21037/atm-20-1789

