J Chin Med Assoc

ORIGINAL ARTICLE

Smartphone-based diabetic macula edema screening with an offline artificial intelligence

De-Kuang Hwang[a,b], Wei-Kuang Yu[a,b], Tai-Chi Lin[a,b], Shih-Jie Chou[b,c], Aliaksandr Yarmishyn[c], Zih-Kai Kao[b,c], Chung-Lan Kao[d,e], Yi-Ping Yang[c,f], Shih-Jen Chen[a,b], Chih-Chien Hsu[a,b*], Ying-Chun Jheng[c,d,e,*]

[a]Department of Ophthalmology, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [b]Faculty of Medicine, National Yang-Ming University, Taipei, Taiwan, ROC; [c]Department of Medical Research, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [d]Department of Physical Medicine & Rehabilitation, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; [e]Department of Physical Medicine and Rehabilitation, School of Medicine, National Yang-Ming University, Taipei, Taiwan, ROC; [f]Department of Pharmacology, Taipei Veterans General Hospital, Taipei, Taiwan, ROC

Abstract

Background: Diabetic macular edema (DME) is a sight-threatening condition that needs regular examinations and remedies. Optical coherence tomography (OCT) is the most common used examination to evaluate the structure and thickness of the macula, but the software in the OCT machine does not tell the clinicians whether DME exists directly. Recently, artificial intelligence (AI) is expected to aid in diagnosis generation and therapy selection. We thus develop a smartphone-based offline AI system that provides diagnostic suggestions and medical strategies through analyzing OCT images from diabetic patients at the risk of developing DME.

Methods: DME patients receiving treatments in 2017 at Taipei Veterans General Hospital were included in this study. We retrospectively collected the OCT images of these patients from January 2008 to July 2018. We established the AI model based on MobileNet architecture to classify the OCT images conditions. The confusion matrix has been applied to present the performance of the trained AI model.

Results: Based on the convolutional neural network with the MobileNet model, our AI system achieved a high DME diagnostic accuracy of 90.02%, which is comparable to other AI systems such as InceptionV3 and VGG16. We further developed a mobile-application based on this AI model available at https://aicl.ddns.net/DME.apk.

Conclusion: We successful integrated an AI model into the mobile device to provide an offline method to provide the diagnosis for quickly screening the risk of developing DME. With the offline property, our model could help those nonophthalmological healthcare providers in offshore islands or underdeveloped countries.

Keywords: Artificial intelligence; Diabetic macular edema; Optical coherence tomography; Smartphone

1. INTRODUCTION

Based on the rapid development of Graphics processing unit (GPU), artificial intelligence (AI) can discriminate a large number of images through a process called deep learning. Current AI models depend on the convolutional neural networks (CNNs) architecture to improve and optimize the deep learning process. AI technique is expected to help diagnosis generation, therapy selection, risk prediction, and disease stratification.[1] In ophthalmology, based on analyzing optical coherence tomography (OCT) images, several studies successfully used AI to detect the existence of single disease manifestation such as the existence of intraretinal fluid, the existence of drusen, or the quantification of macular fluid.[2–4] One possible AI application in this field is to provide screening and diagnostic aid for patients who live in area which is in a lack of ophthalmologists or well-trained optometrists. However, modern networks contain millions of learned connections. The general trend is to design deeper and more complicated networks to achieve higher accuracy. These AI programs usually require high-tech and expensive computer systems containing advanced graphic processing units which are usually unaffordable for utilities in healthcare deficient or low-income areas. In such situation, a mobile smartphone-based AI system with high accuracy and less equipment requirement is extremely important and helpful.

Smartphone applications (app) and mobile robots usually require only low memory and energy footprint.[5] An efficient network architecture, MobileNet, was thus developed to meet the designed requirements of mobile and embedded vision applications. The smaller and faster model uses width multiplier and resolution multiplier by trading off a reasonable amount of accuracy to reduce size and latency. Comparing with the other models, program with MobileNets demonstrated superior size, speed, and accuracy characteristics.[6]

*Address Correspondence: Dr. Chih-Chien Hsu, Department of Ophthalmology, Taipei Veterans General Hospital, 201, Section 2, Shi-Pai Road, Taipei 112, Taiwan, ROC. E-mail address: chihchienym@gmail.com (C.-C. Hsu); Dr. Ying-Chun Jheng, Department of Medical Research, Taipei Veterans General Hospital, 201, Section 2, Shi-Pai Road, Taipei 112, Taiwan, ROC. E-mail address: cycom1220@gmail.com (Y.-C. Jheng)

Conflicts of interest: The authors declare that they have no conflicts of interest related to the subject matter or materials discussed in this article.

Journal of Chinese Medical Association. (2020) 83: 1102-1106.

Received May 8, 2020; accepted May 8, 2020.

doi: 10.1097/JCMA.0000000000000355.

Diabetic macular edema (DME) is the leading cause of blindness in working-age adults.[7,8] Focal laser photocoagulation has been the standard of care to manage DME because the landmark Early Treatment Diabetic Retinopathy Study demonstrated reduction in severe vision loss in patients who received such kinds of treatment.[9] Recently, intravitreal injections (IVI) of antivascular endothelial growth factor (anti-VEGF) agents have progressively replaced focal laser photocoagulation as the primary treatment for center involving macular edema because they provide better vision prognosis for these patients.[8,10] However, regularly monitoring the condition of DME with OCT machine is needed. Because of the rising frequency of obesity, increasing life span, and improved detection of the disease, the number of persons with diabetes worldwide is predicted to grow to 429 million by 2030.[7,11] Using OCT images to determine whether there is DME will be a big burden for ophthalmologists. Previously, emerging studies have shown successful results in training AI systems to identify DME through analyzing OCT images.[12–14] In this study, we are aiming to establish a smartphone-based AI diagnostic platform for DME and to develop a mobile application that is able to provide an offline DME screening by analyzing the uploaded OCT images as illustrated in Fig. 1A.

2. METHODS

The details of the process of setting the AI model are as follows.

2.1. Data collection

DME patients receiving IVI of either anti-VEGF or corticosteroid in 2017 at Taipei Veterans General Hospital were included in this study. We retrospectively collected the OCT images of both eyes from these patients from January 2008 to July 2018. The Institutional Review Board of Taipei Veterans Hospital, Taiwan approved the protocols used in this study.

2.2. Image Preprocessing

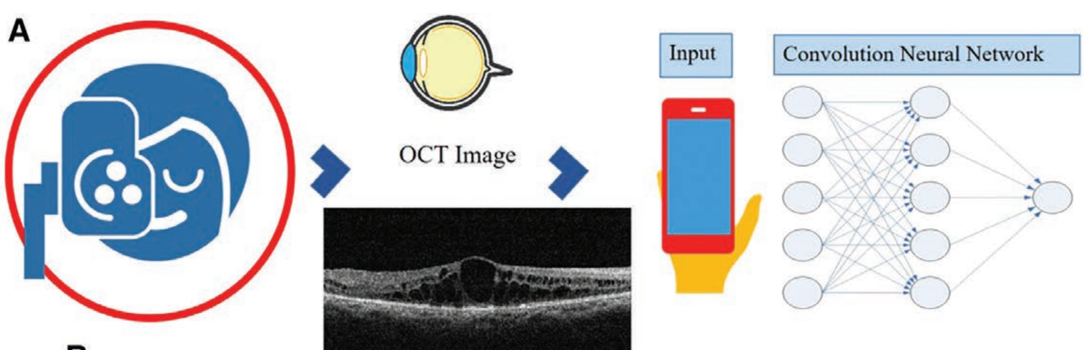
The OCT images were collected from two OCT devices, Zeiss Cirrus HD-OCT 4000 and Optovue RTVue-XR Avanti. Initially, we filtered out OCT images with low resolution or improper format. All enrolled OCT images were 3499 × 2329, 2474 × 2777, or 948 × 879 raw image formats. Resolutions of all images were normalized using the equation of P'i = (Pi – Pmean) / Pstd, in which P'i was the adjusted pixel of each image, Pi denotes each pixel, and Pmean and Pstd are the mean and standard deviation of all pixels. Then, the OCT images were randomly divided into three datasets based on different patients: 80% of the images formed a training dataset, 10% of the images formed a validation dataset, and the rest formed a verification dataset. The OCT images in the training and validation datasets were used to establish AI models, and the verification dataset was used to verify the performance of AI models

2.3. Human image labeling

A senior vitreoretinal surgeon and an experienced ophthalmologist performed DME labeling on the collected OCT images. If there are differences between the labels, another senior retina specialist should be consulted and the three doctors work together to determine the final label. After manual labeling, the identifiable information of the OCT images is removed and all the OCT images are divided into two types: DME and non-DME.

2.4. Dataset augmentation

To overcome the limitation of data size and enhance the performance of the AI model, the images in training dataset have undergone data augmentation in each epoch. The augmented method is not used in validation or verification dataset but only in the training dataset. The following augmentation parameters have been applied and the details of these parameters have been randomly assigned by AI: (1) sheared the image with random angle degrees; (2) image has been zoom in or zoom out



B

Architecture	MobileNet
Optimizer	RMSprop
Learning Rate	1e-4
Batch size	16
Total number of epochs run during training	100
Accuracy of validation of the best model	93.70%

Fig. 1 Schematic illustration and validation accuracy of smartphone artifcial intelligence-based decision-making for diabetic macular edema. A, The macular optical coherence tomography images were obtained from the patients and uploaded to a smartphone. The system was developed based on the convolution neural network and installed in the smartphone as an application. User can get an immediate report of suggestion within few seconds. B, The final model has been trained by RMSprop optimizer with learning rate set as 1e-4, batch size with 16 and the total epoch number was 100. The accuracy of validation dataset of the final model of MobileNet was 93.70%.

randomly; (3) rotate the image with random angle degree; and (4) flip the image with the vertical axis randomly

2.5. Establishment and validation of AI models

Convolution neural network (CNN) has been applied in this study as the main concept to execute the image classification task, the MobileNet has been chosen as the backbone to extract the features of image, and the transfer learning has been applied to enhance the performance of the training task. The classification result of the AI model was represented as a binary value (0 or 1). The loss function and the optimizer function was set as the Sigmoid Cross-Entropy and Root Mean Square Propogation (RMSprop), respectively. Moreover, the input batch size, training epoch, and model learning rate were set as the 16, 100, and 1e-4, respectively. The Google cloud platform has been used to utility our AI model, the main hardware of the platform was NVIDIA Tesla K80 GPU card and 7.5GB RAM. The CentOS7 with Keras 2.2.4 and tensorflow-gpu 1.6.0 was selected as the software to develop the model.

2.6. Verification of AI models and statistical analysis

Accuracy, specificity, and sensitivity have been used to evaluate the performance of our trained AI model. Based on the concept of the confusion matrix, those parameters has been defined as the following equations:

$$\text{Accuracy} = (TP + TN)/(TP + FP + TN + FN), \quad (1)$$

$$\text{Specificity} = TN/(FP + TN), \quad (2)$$

$$\text{Sensitivity} = TP/(TP + FN) \quad (3)$$

where TP, TN, FP, and F*N was indicated the true positive, true negative, false positive, and false negative, respectively.

Receiver operating characteristic (ROC) curve has been applied to represent AI model performance. The horizontal and vertical axes have been defined as the false positive rate (FPR) and true positive rate (TPR), respectively. The TPR meant the sensitivity and the FPR was the result of 1 subtracts the specificity value. Based on the ROC, we calculate the area under the curve (AUC) and use it to assess the performance of our AI model. The value of AUC is between 0.5 and 1, and the higher the value was, the more correct the AI model's predictions.

3. RESULTS

3.1. Image collection

The OCT images were collected from 173 diabetic patients and randomly selected in the database based on different patients. A total number of 4932 OCT images were collected but only 3495 passed the image quality control and were used in this study. Among these, 80% (2768) of images based on different patients were randomly selected as the training dataset to establish our AI models, 365 images were selected as a validation dataset, and 362 images were selected as a verification dataset. The images in training dataset have been executed the data augmentation, which was only applied for perturbing the training dataset, but not for validation or verification dataset.

3.2. Establishing AI models

The CNN architecture, MobileNet, was applied to establish the AI model. There were three main layers, convolution layer, pooling layer, and fully connected layer in the CNN architecture. During the convolution layer, the feature detector will have applied the executed the convolution, to filter the unnecessary information and trace the important features (such as shape or color). After several convolution layers, the size of traced features will become enormous and make burden for computing resource. Hence, the pooling layer will be applied to filter the noise and keep the most important features, which can not only retain the important feature for AI model learning but also accelerate the training process. Finally, the fully connected layer will flatten all of those features, which make the important information from two-dimensional transfer to one-dimensional transfer. According the one-dimensional information, the AI model will make decision for classifying the input image is the DME or non-DME patients and provide the final result for us.

The images were adjusted to the same size, and the skill of data augmentation was used to enhance the efficiency of our model. The batch size of the training layers has been set as 16 images per step and used the RMSprop optimizer with a learning rate of 1e-4. Each model with 100 epochs for training, the more iterative the epoch was, and the more accurate the AI model. The best models with the minimal value of loss were selected for the verification. The accuracy of MobileNet to recognize the validation dataset was 93.70% (Fig. 1B).

3.3. Verification of the final model

The performance of the AI model, MobileNet, has been verified by the validation dataset containing 227 DME and 135 non-DME OCT images. As shown in the confusion matrix (Fig. 2A), the accuracy of the AI model was 90.06%, the sensitivity of the AI model was 92.51%, and the specificity was 85.93%. The ROC curve and the AUC of this AI models was calculated as shown in Fig. 2B. The AUC of MobileNet was 0.96.

3.4. Development of app-based service for DMD diagnostics

We developed a mobile application to provide Smartphone-based service with our AI model, which is available at https://aicl.ddns.net/DME.apk. The app could be downloaded and installed in each smartphone with android system. By this app, no matter where users are, all they have to do is open their OCT images and press the diagnosis button, then the result of diagnosis and medical suggestion will display on the monitor (Fig. 3). The diagnosis normally takes 3 to 6 seconds. According the diagnosis result, the AI model can provide a recommendation for the user.

4. DISCUSSION

In this article, we established an offline smartphone-based AI (MobileNet) screening platform for DME, which can analyze OCT images taken from patients with diabetes mellitus. The diagnostic accuracy of our smartphone-based AI model for detection of DME is 90.02%, which is similar to computer-based AI models, such as VGG16 and InceptionV3. Based on this AI model, we developed a mobile application on Android system, available at https://aicl.ddns.net/DME.apk. To the best of our knowledge, this is the first study evaluating an offline smartphone-based AI screening platform for DME in the form of mobile application.

With the popularity of smartphones and Internet of Things (IoT), AI edge computing was considered to be implemented into smartphones for providing real-time service.[15,16] The advantages of AI edge computing are low latency, stability, reliability, and privacy in the processing and transmission of data, particularly as advantages of offline processing are concerned.[17] With the AI edge computing, the use of limited computing resources to achieve the maximum performance was the most important issue.[18]

The core novelty of MobileNet separated the convolution layer into two parts, depthwise and pointwise.[6] The depthwise

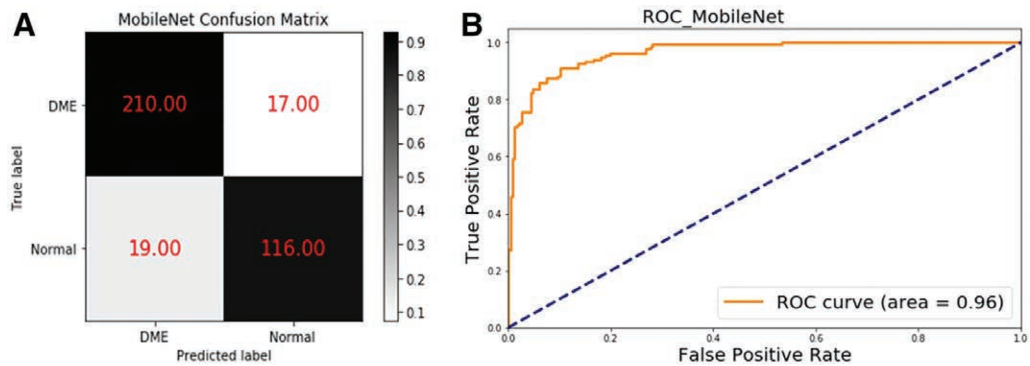


Fig. 2 Confusion matrix and ROC curves of three CNN-based AI model. Confusion matrix and ROC curves of three CNN-based AI model. A, Confusion matrix indicate the accuracy of the MobileNet-based AI model (90.06%). B, AUC of MobileNet presents good performance (0.96). AI = artificial intelligence; AUC = area under the curve; CNN = convolution neural network; ROC = receiver operating characteristic.

Fig. 3 Schematic illustration of smartphone AI-based decision-making for diabetic macular edema. After examination, technicians or users import their optical coherence tomography images into their mobile phone. The AI application immediately provides high-accuracy interpretation and offer user the remedy recommendation. The system could also refer the images with critical signs or undifferentiated results to healthcare providers (trained specialists) who may make the final diagnosis and treatment decision. AI = artificial intelligence.

gathered the features from each channel and the pointwise was the tradition convolution filter, but the size was set as a 1*1 feature detector. In fact, the performance of depthwise layer and the pointwise layers was similar to a standard convolution, but it will greatly reduce the amount of calculation resource and model burden.[19] Besides, the RMSprop has been selected as the optimizer function. This function is beneficial to eliminate the direction of large swing amplitude, and is used to correct the swing amplitude, so that the swing amplitude in each dimension is smaller. On the other hand, it also makes the network function converge faster.

Most of the OCT images could be converted into grayscales and have relatively lower resolutions than other ocular images such as fundus photography. Our study implicated that the AI system might obtain a sufficient ability to recognize some significant difference in picture such as the appearance of intraretinal or subretinal fluid in DME by training with relatively smaller dataset. With the main purpose of creating a mobile and easy-accessible screening and diagnostic-aiding program in this study, we believed that our MobileNet-derived AI program achieved a satisfying performance in identifying DME. Besides, the accuracy of disease classification could possibly be improved with newly developed network architecture after training with a larger dataset.[20]

In conclusion, we integrated the AI into the mobile phone to provide an accessible method for most population and provide the scientific evidence to verify the reliability and feasibility in real world. With the application of mobile phone, patients can upload their image and then obtain accurate diagnosis and recommended medical care whether where they are. The AI-based mobile phone system can help specialists and patients from remote mountainous, offshore islands, or underdeveloped countries. Furthermore, we hope it will also promote the benefits of medical experts, patients, and other related industries. In the future, we plan to integrate the e-cloud technology to establish an AI-based telemedicine remote medical application for the mobile phone system. The system can not only provide medical service for the user but also collect the feedback and image information from the user to enhance the performance and expand the application territories.

ACKNOWLEDGMENTS

This research was funded by the Taiwan Ministry of Science and Technology (MOST-108-2314-B-010-042-MY3, MOST-108-2811-B-010-541, MOST-108-2314-B-075-055), Taipei Veterans General Hospital (CI-109-19). We thank Ying-Hsuan Wu, Hou-Kai Hu, Feng-Yuan Yang in data collection, and thanks Hsin-Yi Huang served as scientific advisors of the manuscript. This study was assisted in part by the Division of Experimental Surgery of the Department of Surgery and the Animal Center of Taipei Veterans General Hospital.

REFERENCES

1. Hwang DK, Hsu CC, Chang KJ, Chao D, Sun CH, Jheng YC, et al. Artificial intelligence-based decision-making for age-related macular degeneration. *Theranostics* 2019;9:232–45.
2. Zhao R, Camino A, Wang J, Hagag AM, Lu Y, Bailey ST, et al. Automated drusen detection in dry age-related macular degeneration by multiple-depth, en face optical coherence tomography. *Biomed Opt Express* 2017;8:5049–64.
3. Lee CS, Tyring AJ, Deruyter NP, Wu Y, Rokem A, Lee AY. Deep-learning based, automated segmentation of macular edema in optical coherence tomography. *Biomed Opt Express* 2017;8:3440–8.

4. Schlegl T, Waldstein SM, Bogunovic H, Endstraßer F, Sadeghipour A, Philip AM, et al. Fully automated detection and quantification of macular fluid in OCT Using deep learning. *Ophthalmology* 2018;**125**:549–58.
5. Tung F, Mori G. Deep neural network compression by in-parallel pruning-quantization. *IEEE Trans Pattern Anal Mach Intell* 2020;**42**:568–79.
6. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. 2017. Available at https://arxiv.org/abs/1704.04861. Accessed April 17, 2017.
7. Antonetti DA, Klein R, Gardner TW. Diabetic retinopathy. *N Engl J Med* 2012;**366**:1227–39.
8. Korobelnik JF, Do DV, Schmidt-Erfurth U, Boyer DS, Holz FG, Heier JS, et al. Intravitreal aflibercept for diabetic macular edema. *Ophthalmology* 2014;**121**:2247–54.
9. ETDRS research group. Photocoagulation for diabetic macular edema. Early treatment diabetic retinopathy study report number 1. Early treatment diabetic retinopathy study research group. *Arch Ophthalmol* 1985;**103**:1796–806.
10. Nguyen QD, Brown DM, Marcus DM, Boyer DS, Patel S, Feiner L, et al.; RISE and RIDE Research Group. Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. *Ophthalmology* 2012;**119**:789–801.
11. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004;**27**:1047–53.
12. Alsaih K, Lemaitre G, Rastgoo M, Massich J, Sidibé D, Meriaudeau F. Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. *Biomed Eng Online* 2017;**16**:68.
13. Gerendas BS, Bogunovic H, Sadeghipour A, Schlegl T, Langs G, Waldstein SM, et al. Computational image analysis for prognosis determination in DME. *Vision Res* 2017;**139**:204–10.
14. Prahs P, Radeck V, Mayer C, Cvetkov Y, Cvetkova N, Helbig H, et al. OCT-based deep learning algorithm for the evaluation of treatment indication with anti-vascular endothelial growth factor medications. *Graefes Arch Clin Exp Ophthalmol* 2018;**256**:91–8.
15. Ai Y, Peng M, Zhang K. Edge computing technologies for internet of things: a primer. *DCN* 2018;**4**:77–86.
16. Calo SB, Touna M, Verma DC, Cullen A. Edge computing architecture for applying AI to IoT. In: 2017 IEEE International Conference on Big Data (Big Data), 11–14 December. 2017. Doi:10.1109/BigData.2017.8258272.
17. Lin J, Yu W, Zhang N, Yang X, Zhang H, Zhao W. A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications. *IEEE Internet Things J* 2017;**4**:1125–42.
18. Wang X, Han Y, Wang C, Zhao Q, Chen X, Chen M. In-edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning. *IEEE Network* 2019;**33**:156–65.
19. Singh B, Toshniwal D, Allur SK. Shunt connection: An intelligent skipping of contiguous blocks for optimizing MobileNet-V2. *Neural Netw* 2019;**118**:192–203.
20. Ristevski B, Chen M. Big data analytics in medicine and healthcare. *J Integr Bioinform* 2018;**15**:20170030.