Original Article

Detection and Classification of Diabetic Macular Edema with a Desktop-Based Code-Free Machine Learning Tool

Furkan Kırık, Büşra Demirkıran, Cansu Ekinci Aslanoğlu, Arif Koytak, Hakan Özdemir

Bezmialem Vakif University Faculty of Medicine, Department of Ophthalmology, İstanbul, Türkiye

Abstract

Objectives: To evaluate the effectiveness of the Lobe application, a machine learning (ML) tool that can be used on a personal computer without requiring coding expertise, in the recognition and classification of diabetic macular edema (DME) in spectral-domain optical coherence tomography (SD-OCT) scans.

Materials and Methods: A total of 695 cross-sectional SD-OCT images from 336 patients with DME and 200 OCT images of 200 healthy controls were included. Images with DME were classified into three main types: diffuse retinal edema (DRE), cystoid macular edema (CME), and cystoid macular degeneration (CMD). To develop the ML model, we used the desktop-based code-free Lobe application, which includes a pre-trained ResNet-50 V2 convolutional neural network and is available free of charge. The performance of the trained model in recognizing and classifying DME was evaluated with 41 DRE, 28 CMD, 70 CME, and 40 normal SD-OCT images that were not used in the training.

Results: The developed model showed 99.28% sensitivity and 100% specificity for class-independent detection of DME. Sensitivity and specificity by labels were 87.80% and 98.57% for DRE, 96.43% and 99.29% for CME, and 95.71% and 95.41% for CMD, respectively.

Conclusion: To our knowledge, this is the first evaluation of the effectiveness of Lobe with ophthalmological images, and the results indicate that it can be used with high efficiency in the recognition and classification of DME from SD-OCT images by ophthalmologists without coding expertise.

Keywords: Artificial intelligence, code-free, machine learning, diabetic macular edema, optical coherence tomography

Introduction

Optical coherence tomography (OCT) provides physicians with non-invasive, rapid, and micron-level resolution images of the ocular tissues with near histological detail. It is widely used in diagnosis and follow-up of many retinal diseases, especially diabetic retinopathy.[1] Diabetic macular edema (DME), which is the main cause of visual impairment in patients with diabetic retinopathy, can also be successfully detected with OCT.[2] Various OCT-based DME classifications have been developed.[3,4,5,6,7] The most current of these classifications was reported by Arf et al.[7] and defines three main types of DME with different clinical and morphological features: diffuse retinal edema (DRE), cystoid macular edema (CME), and cystoid macular degeneration (CMD).

OCT has become more preferred for the development of artificial intelligence (AI) models compared to other imaging methods due to its widespread usage and superiority in terms of acquiring multiple retinal images, providing high-resolution retinal imaging of pathological lesions undetectable by standard color fundus photography or clinical examination, and revealing various biomarkers that can provide information about disease prognosis.[8,9] Although physicians' interest in AI has increased over time, many still have reservations about AI because earlier systems required a certain level of coding skills, and highly specialized computing resources were needed.[10]

With the web-, cloud-, or personal computer-based code-free AI platforms that have become available in recent years, physicians can develop their own AI models and perform classification and segmentation of medical images without the need for any coding expertise.[11] Lobe (www.lobe.ai, Lobe Artificial Intelligence, Microsoft, Inc.) is a free desktop-based no-code machine learning (ML) application that classifies images by using pre-trained ResNet-50 V2 and MobileNet V2 convolutional neural networks (CNN).[12] However, the effectiveness of Lobe in classifying ocular images is unknown.

The aim of this study was to evaluate the effectiveness of Lobe in the detection and classification of DME from cross-sectional spectral-domain (SD)-OCT scans.

Materials and Methods

Dataset Preparation and Image Labeling

Macular volumetric OCT scans of 336 patients with diabetic retinopathy and DME detected by Heidelberg Spectralis SD-OCT (Heidelberg Engineering, Inc., Heidelberg, Germany) between June 2019 and June 2021 were retrospectively analyzed. Cross-sectional OCT scans were evaluated by two ophthalmologists (H.O., F.K.) according to DME type, image quality, and the presence of additional retinal pathology. The classification defined by Arf et al.[7] was used to determine DME type. Accordingly, DRE was defined as DME characterized by increased retinal thickness and decreased intraretinal reflectivity, without a prominent round or oval intraretinal fluid space (Figure 1a). CME was defined as DME containing hyporeflective round or oval-shaped intraretinal cystoid areas bordered by hyperreflective septa (Figure 1b). CMD was defined as DME containing 600 µm and larger intraretinal cystoid spaces (Figure 1c). In the Arf et al.[7] classification, the presence of serous macular detachment, vitreomacular interface disease, and hard exudate is designated as subgroups a, b, and c, respectively. Although there was no image subclassification based on the presence of these pathological lesions in this study, images with these findings were also included. Eyes with additional retinal pathology, such as age-related macular degeneration, glaucoma, and images with signal quality worse than 20 (according to manufacturer's signal quality index, range 0-40) were excluded.

As a result, a total of 695 fovea-centered cross-sectional SD-OCT images were included and 205 were classified as DRE, 350 as CME, and 140 as CMD-type DME. In addition, 200 fovea-centered SD-OCT scans from 200 healthy controls were included as "normal" (Figure 1d). All images were resized to 512 x 512 pixels centered on the fovea using the crop function in Fiji (ImageJ, 1.53f; National Institute of Health, Bethesda, MD, USA). Metadata such as patient and device information were deleted from the images. Eighty percent of the images were parsed for training the ML model and 20% were allocated for testing.

This study has been approved by the Bezmialem Vakif University Faculty of Medicine Ethics Committee (decision no: 2022/30, date: 22.02.2022).

Training the Deep Learning Model

The Lobe program was downloaded from the website https://www.lobe.ai/ (Version 0.10.1130.5) free of charge and installed on a personal computer. After installation, a new project was created with four image labels named DRE, CMD, CME, and normal (Figure 2). To train the ML model, 164 DRE, 280 CME, and 112 CMD, and 160 normal images were imported into the appropriate image classes created in the application.

No additional data augmentation techniques were performed, as the current application automatically generated five random variations of the image (random modifications included brightness, contrast, saturation, hue, rotation, zoom, and JPEG encoding noise) during training.[12] The ResNet-50 V2 CNN was used for model architecture by selecting the "optimize for accuracy" option in the Project Settings menu. After selecting the CNN, the training phase was restarted, followed by model optimization with the built-in "model optimization" function for better real-world performance.

Performance Evaluation and Statistical Analysis

Twenty percent of the dataset, consisting of 41 DRE, 28 CMD, 70 CME, and 40 normal external test images allocated for testing, were imported one by one into the "Use" function of Lobe, and the model's prediction for each image was recorded (Figure 3). Predictions were not manually marked as "correct" or "incorrect" during testing.

Statistical analyses were performed using SPSS version 22 software package (IBM Corp., Armonk, NY, USA). The sensitivity and specificity of the model in recognizing any type of DME (DRE or CME or CMD vs. normal image) and in detecting individual types of DME were determined. In addition, the area under the curve (AUC) was calculated with receiver operating characteristic (ROC) curve analysis to determine the effectiveness of the model in image classification.
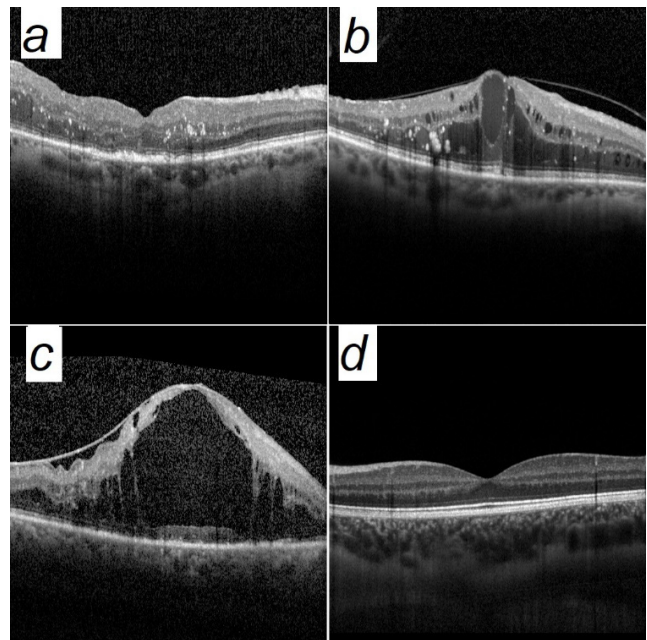

Figure 1. Samples of retinal spectral-domain optical coherence tomography (SD-OCT) scans from different classes included in the study. SD-OCT images of diabetic macular edema were classified into three types: diffuse retinal edema, defined as retinal thickening with reduced retinal hyperreflectivity and no cystoid spaces (a); cystoid macular edema, defined as retinal thickening with round or oval intraretinal spaces with a diameter less than 600 microns separated by hyperreflective septa (b); and cystoid macular degeneration, defined as intraretinal hyporeflective cystoid spaces with a diameter greater than 600 microns (c). SD-OCT images with no retinal pathology were classified as "normal" (d)

Results

Internal validation automatically performed using 145 internal images randomly selected from the imported images used for model training indicated that the developed model had prediction accuracy of 93.79% (98.74% of all imported images) in labeling. The effectiveness of the model was also evaluated with 41 DRE, 28 CMD, 70 CME, and 40 normal external test images. The predictions of the ML model in classifying DME types on external test images and the normalized confusion matrix based on the test data are shown in Figure 4.

Sensitivity, specificity, and AUC values for detecting DME regardless of type (any type of DME vs. normal) were calculated as 99.28%, 100%, and 0.996, respectively. Sensitivity, specificity, and AUC values for each DME type were calculated as 87.80%, 98.57%, and 0.936 for DRE; 96.43%, 99.29%, and 0.979 for CMD; and 95.71%, 95.41%, and 0.960 for CME, respectively. ROC curve analyses indicating the performance of the ML model in classifying DME types are shown in Figure 5. The presence of serous macular detachment, vitreomacular interface disease, and hard exudates in the training and testing datasets are also presented in Table 1.

Discussion

This study demonstrated that Lobe, a code-free ML application, could be used with high efficiency in the diagnosis and classification of DME from cross-sectional SD-OCT images. In addition, it showed that Lobe could provide satisfactory performance with a much smaller number of images compared to untrained deep learning (DL) models. This is an advantage of performing transfer learning using a pre-trained algorithm and automatically applying data augmentation to the dataset.

There are several studies in which ophthalmological images were classified by computer vision, a subfield of AI.[8,9,13,14,15,16,17,18,19] However, few studies have attempted to classify DME subtypes from OCT images. Alsaih et al.[18] developed a multi-stage ML model to identify the presence of retinal thickening (DRE in the new classification), hard exudates, intraretinal cystoid spaces (CME in the new classification), and subretinal fluid (serous macular detachment in new classification) using the volumetric SD-OCT scans of 16 patients. The generic pipeline they developed included pre-processing, feature detection, feature representation, and classification. Although this model appears to have a sequential structure that can perform the current DME classification, it may not be practical for physicians without prior coding knowledge. In addition, it differs from our study in that all images featuring diabetic retinopathy were obtained from only 16 patients and classic ML algorithms were used instead of the CNN that forms the structure in Lobe.

A method similar to that performed in our study was recently used by Wu et al.[19] The authors aimed to classify DRE, CMD, and serous macular detachment based on SD-OCT images using a VGG-16 CNN. Their model was developed with a large number of OCT images (12365 in total), yet shows little superiority over the DRE and CME classification in our study (AUC values 0.970 vs. 0.936 for DRE and 0.997 vs. 0.960 for CME, respectively). A large amount of data is required, especially during the development of DL models, and therefore the number of OCT images included in our study may seem insufficient.[20,21] Despite this, the DL model developed in our study resulted in acceptable accuracy with a relatively small number of images. The main reason for this may be the fact that Lobe utilizes transfer learning with pre-trained weights from the ImageNet, and automatically applies data augmentation techniques to the dataset. Thus, it was demonstrated that Lobe trained with a small number of
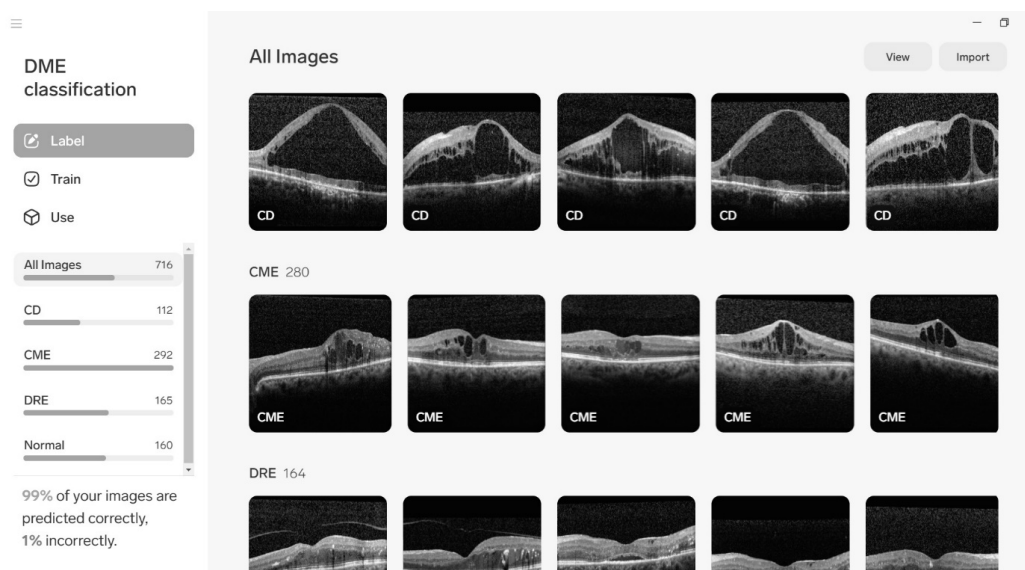

Figure 2. The "New Project" user interface (named "DME classification" in the current project) in the Lobe program. Spectral-domain optical coherence tomography images were labeled as diffuse retinal edema (DRE), cystoid macular edema (CME), cystoid macular degeneration (CD), and normal according to the classes defined after being imported
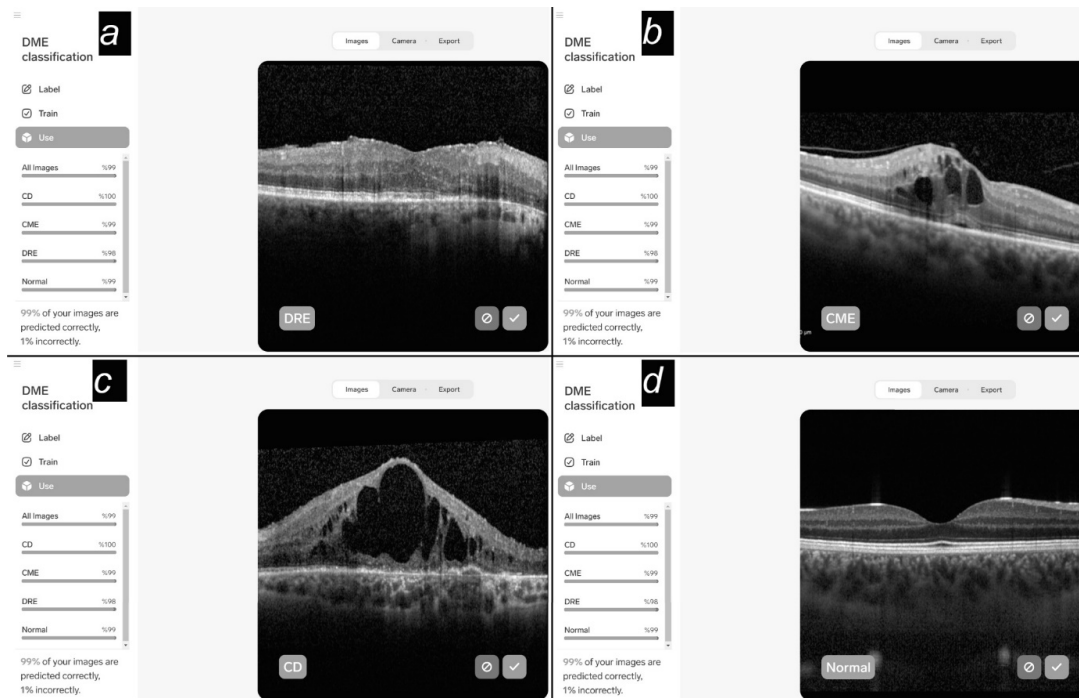
Figure 3. The accuracy of the developed model in predicting the labels was evaluated on external test images by the "Use" function. Model prediction on the sample external test images: (a) diffuse retinal edema (DRE), (b) cystoid macular edema (CME), (c) cystoid macular degeneration (labeled CD in the project), and (d) normal images
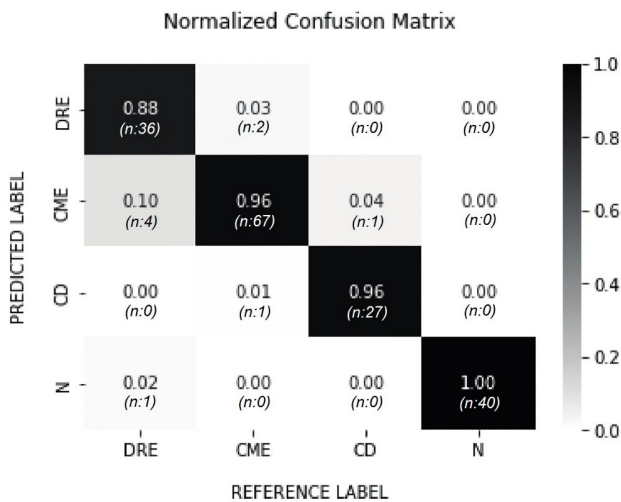

Figure 4. Normalized confusion matrix graph indicates the model had high accuracy in the prediction of diffuse retinal edema (DRE), cystoid macular edema (CME), cystoid macular degeneration (labeled CD in the project), and normal images

OCT images could be used effectively in DME classification and recognition, similar to DL-based models developed using large datasets. Furthermore, the built-in automatic data augmentation function in Lobe seems to be an added benefit, without the need for additional software/coding for data augmentation.

With increasing interest in AI and its widespread use in recent years, many companies have started to offer no-code ML platforms to serve users who do not have coding experience. The most detailed analysis of these platforms, which have different features and functions, was conducted recently by Korot et al.[11] Their study examined the performance of various code-free DL platforms in open-access datasets, as well as several features such as data security, usage fee, and model architecture. However, Lobe was not evaluated in that study. To the best of our knowledge, the clinical usage of Lobe has recently been evaluated in a few non-ophthalmological studies but not in the field of ophthalmology.[22] The cost-free availability of Lobe is an advantage for physicians seeking to gain experience in AI with image classification. In addition, running and training DL models entirely on a personal device, i.e., not having to share data with the cloud or web services, seems to be a feature that can satisfy users in terms of data security. Another remarkable feature of Lobe is that after the model is developed, it can be improved even while in use. The user can confirm the prediction made by the model for an uploaded image as correct or incorrect during the test phase of the model. Upon selecting "correct" or "reject" in the "use" tool, the image is saved automatically with the label that the user deems appropriate, allowing the model to improve after each test image.

Lobe includes two different CNN architectures and is selected automatically based on the size and complexity of the dataset, or manually based on user preference.[12] In this study, we preferred the ResNet-50 V2 CNN structure because we targeted high accuracy from the model. On the other hand, the user can also choose to use the MobileNet V2 to get
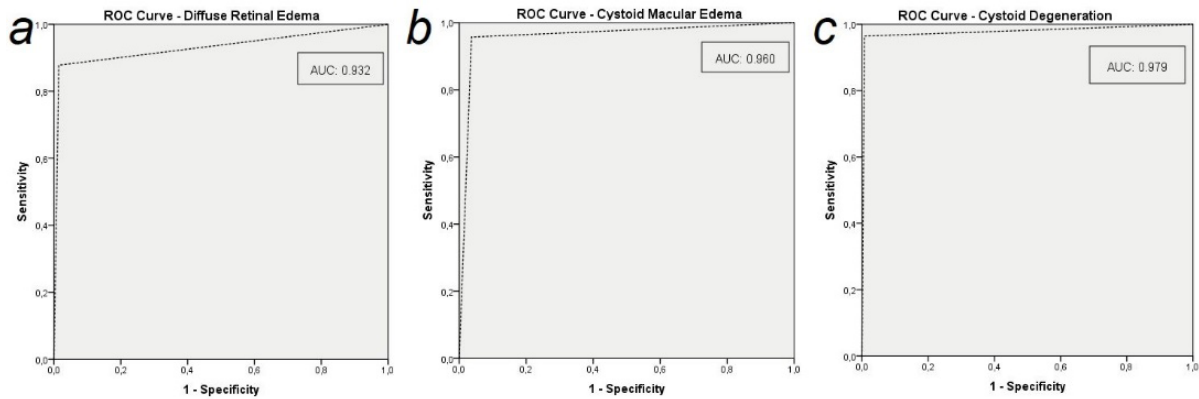
Figure 5. Receiver operating characteristic (ROC) curves and area under the ROC curves (AUC) indicate the diagnostic ability of the developed model for the types of diabetic macular edema: diffuse retinal edema (a), cystoid macular edema (b), and cystoid macular degeneration (c)

Table 1. The presence of serous macular detachment, vitreomacular interface disease, and hard exudate in the training and test datasets

	DRE		CME		CMD	
	Train (n=164)	Test (n=41)	Train (n=280)	Test (n=70)	Train (n=112)	Test (n=28)
SMD (%)	46 (28.05)	7 (17.07)	99 (35.36)	18 (25.71)	5 (4.46)	1 (3.57)
VMID (%)	31 (18.90)	9 (21.95)	20 (7.14)	4 (5.71)	19 (16.96)	6 (21.43)
HE (%)	75 (45.73)	19 (46.34)	138 (49.29)	17 (24.29)	19 (16.96)	3 (10.71)
SMD: Serous macular detachment, VMID: Vitreomacular interface disease, HE: Hard exudate, DRE: Diffuse retinal edema, CME: Cystoid macular edema, CMD: Cystoid macular degeneration						

faster predictions at the cost of lower accuracy.[23,24] In addition, these developed models can be exported and used as a mobile application if desired. Although Lobe has many advantageous features and can successfully identify and classify DME in SD-OCT images, it lacks object detection or segmentation functions that would allow the identification of target structures in images, which limits its areas of application.

Study Limitations

The difference in the number of images between groups is a limitation of this study. However, in our retrospectively collected body of data, images with CME were frequently encountered, while images with CMD were less common. As reported by Arf et al.,[7] the frequency of CMD was lower than other types of DME. In addition, a larger data sample in DL/ML models increases the accuracy of the models. However, imbalance (over- or under-sampling) in datasets is a problem that limits the accuracy and reliability of the models. A size ratio of the smallest (minority) class (group) to the largest (majority) class (group) between 20% and 40% is considered a mild imbalance.[25] For this reason, datasets were created to attain high accuracy from the model without being imbalanced. Similar sample size differences between the groups are also seen in previous studies.[26]

Conclusion

Our study showed that Lobe, a free, no-code ML program, could be used effectively in DME detection and classification without the need for a large dataset by using pre-trained CNN architecture and automatic data augmentation. Advantageous features of the program are that it provides additional data security by being used on a personal device, the developed model can continue to be improved with every test image, and users have the option of selecting the CNN architecture. Thus, the Lobe program is an efficient and user-friendly option for physicians who do not have basic coding skills.

Ethics

Ethics Committee Approval: This study has been approved by the Bezmialem Vakif University Faculty of Medicine Ethics Committee (decision no: 2022/30, date: 22.02.2022).

Informed Consent: Obtained.

Peer-review: Externally peer-reviewed.

Authorship Contributions

Surgical and Medical Practices: H.Ö., A.K., Concept: F.K., H.Ö., A.K., Design: F.K., H.Ö., Data Collection or Processing: B.D., C.E.A., Analysis or Interpretation: H.Ö., A.K., Literature Search: F.K., B.D., C.E.A., Writing: F.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Fujimoto JG, Drexler W. Introduction to OCT. In: Drexler W, Fujimoto JG, eds. Optical Coherence Tomography: Technology and Applications. 2nd ed. Cham; Springer; 2015:3-64.

2. Acón D, Wu L. Multimodal Imaging in Diabetic Macular Edema. Asia Pac J Ophthalmol (Phila). 2018;7:22-27.

3. Ruia S, Saxena S, Gemmy Cheung CM, Gilhotra JS, Lai TY. Spectral Domain Optical Coherence Tomography Features and Classification Systems for Diabetic Macular Edema: A Review. Asia Pac J Ophthalmol (Phila). 2016;5:360-367.

4. Otani T, Kishi S, Maruyama Y. Patterns of diabetic macular edema with optical coherence tomography. Am J Ophthalmol. 1999;127:688-693.

5. Panozzo G, Parolini B, Gusson E, Mercanti A, Pinackatt S, Bertoldo G, Pignatto S. Diabetic macular edema: an OCT-based classification. Semin Ophthalmol. 2004;19:13-20.

6. Kim BY, Smith SD, Kaiser PK. Optical coherence tomographic patterns of diabetic macular edema. Am J Ophthalmol. 2006;142:405-412.

7. Arf S, Sayman Muslubas I, Hocaoglu M, Ersoz MG, Ozdemir H, Karacorlu M. Spectral domain optical coherence tomography classification of diabetic macular edema: a new proposal to clinical practice. Graefes Arch Clin Exp Ophthalmol. 2020;258:1165-1172.

8. Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, Tan GSW, Schmetterer L, Keane PA, Wong TY. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019;103:167-175.

9. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina. Prog Retin Eye Res. 2018;67:1-29.

10. Faes L, Wagner SK, Fu DJ, Liu X, Korot E, Ledsam JR, Back T, Chopra R, Pontikos N, Kern C, Moraes G, Schmid MK, Sim D, Balaskas K, Bachmann LM, Denniston AK, Keane PA. Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. Lancet Digit Health. 2019;1:232-242.

11. Korot E, Guan Z, Ferraz D, Wagner SK, Zhang G, Liu X, Faes L, Pontikos N, Finlayson SG, Khalid H, Moraes G, Balaskas K, Denniston AK, Keane PA. Code-free deep learning for multi-modality medical image classification. Nat Mach Intell. 2021;3:288-298.

12. Lobe Help: What models are used? Available from: https://www.lobe.ai/docs/train/train. Access date: 01 May 2022

13. Hormel TT, Hwang TS, Bailey ST, Wilson DJ, Huang D, Jia Y. Artificial intelligence in OCT angiography. Prog Retin Eye Res. 2021;85:100965.

14. Ting DSJ, Foo VH, Yang LWY, Sia JT, Ang M, Lin H, Chodosh J, Mehta JS, Ting DSW. Artificial intelligence for anterior segment diseases: Emerging applications in ophthalmology. Br J Ophthalmol. 2021;105:158-168.

15. Li JO, Liu H, Ting DSJ, Jeon S, Chan RVP, Kim JE, Sim DA, Thomas PBM, Lin H, Chen Y, Sakomoto T, Loewenstein A, Lam DSC, Pasquale LR, Wong TY, Lam LA, Ting DSW. Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective. Prog Retin Eye Res. 2021;82:100900.

16. Reid JE, Eaton E. Artificial intelligence for pediatric ophthalmology. Curr Opin Ophthalmol. 2019;30:337-346.

17. Keskinbora K, Güven F. Artificial Intelligence and Ophthalmology. Turk J Ophthalmol. 2020;50:37-43.

18. 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.

19. Wu Q, Zhang B, Hu Y, Liu B, Cao D, Yang D, et al. Detection of Morphologic Patterns of Diabetic Macular Edema Using A Deep Learning Approach Based on Optical Coherence Tomography Images. Retina. 2021;41:1110-1117.

20. Ting DSW, Lee AY, Wong TY. An Ophthalmologist's Guide to Deciphering Studies in Artificial Intelligence. Ophthalmology. 2019;126:1475-1479.

21. Soekhoe D, van der Putten P, Plaat A. On the Impact of Data Set Size in Transfer Learning Using Deep Neural Networks. In: Boström H, Knobbe A, Soares C, Papapetrou P, eds. Advances in Intelligent Data Analysis XV. Cham; Springer; 2010:50-60.

22. Richardson ML, Ojeda PI. A "Bumper-Car" Curriculum for Teaching Deep Learning to Radiology Residents. Acad Radiol. 2022;29:763-770.

23. Keras Applications. Available from: https://keras.io/api/applications/. Access date: 01 May 2022.

24. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. Mobilenetv2: Inverted residuals and linear bottlenecks. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018:4510-4520.

25. Imbalanced Data. Available from: https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data?hl=en. Access date: 31 March 2023.

26. Milea D, Najjar RP, Zhubo J, Ting D, Vasseneix C, Xu X, Aghsaei Fard M, Fonseca P, Vanikieti K, Lagrèze WA, La Morgia C, Cheung CY, Hamann S, Chiquet C, Sanda N, Yang H, Mejico LJ, Rougier M-B, Kho R, Thi Ha Chau T, Singhal S, Gohier P, Clermont-Vignal C, Cheng C-Y, Jonas JB, Yu-Wai-Man P, Fraser CL, Chen JJ, Ambika S, Miller NR, Liu Y, Newman NJ, Wong TY, Biousse V; BONSAI Group. Artificial Intelligence to Detect Papilledema from Ocular Fundus Photographs. N Engl J Med. 2020;382:1687-1695.