Int Ophthalmol (2019) 39:1871–1877
https://doi.org/10.1007/s10792-018-1016-x

ORIGINAL PAPER

Comparison between support vector machine and deep learning, machine-learning technologies for detecting epiretinal membrane using 3D-OCT

Tomoaki Sonobe · Hitoshi Tabuchi · Hideharu Ohsugi · Hiroki Masumoto · Naohumi Ishitobi · Shoji Morita · Hiroki Enno · Daisuke Nagasato

Received: 6 April 2018 / Accepted: 1 September 2018 / Published online: 14 September 2018

Abstract

Purpose In this study, we compared deep learning (DL) with support vector machine (SVM), both of which use three-dimensional optical coherence tomography (3D-OCT) images for detecting epiretinal membrane (ERM).

Methods In total, 529 3D-OCT images from the Tsukazaki hospital ophthalmology database (184 non-ERM subjects and 205 ERM patients) were assessed; 80% of the images were divided for training, and 20% for test as follows: 423 training (non-ERM 245, ERM 178) and 106 test (non-ERM 59, ERM 47) images. Using the 423 training images, a model was created with deep convolutional neural network and SVM, and the test data were evaluated.

Results The DL model's sensitivity was 97.6% [95% confidence interval (CI), 87.7–99.9%] and specificity was 98.0% (95% CI, 89.7–99.9%), and the area under the curve (AUC) was 0.993 (95% CI, 0.993–0.994). In contrast, the SVM model's sensitivity was 97.6% (95% CI, 87.7–99.9%), specificity was 94.2% (95% CI, 84.0–98.7%), and AUC was 0.988 (95% CI, 0.987–0.988).

Conclusion DL model is better than SVM model in detecting ERM by using 3D-OCT images.

Keywords Deep learning · Support vector machine · Epiretinal membrane · Optical coherence tomography

T. Sonobe (✉) · H. Tabuchi · H. Ohsugi · H. Masumoto · N. Ishitobi · D. Nagasato
Department of Ophthalmology, Tsukazaki Hospital, 68-1 Waku, Aboshi-ku, Himeji 671-1227, Japan
e-mail: t.sonobe@tsukazaki-eye.net

S. Morita
Research Group of Intelligent Cybernetics and Computer Science Graduate School of Engineering, University of Hyogo, Himeji, Japan

H. Enno
Rist Inc, Tokyo, Japan

Introduction

Epiretinal membrane (ERM) is the development of avascular fibrous proliferating tissue on the inner limiting membrane of the upper layer of the retina [1]. Tangential retraction of the retina occurs due to various changes in the eye caused by numerous conditions including cortical remnants by posterior vitreous detachment, retinal tear, vascular occlusion and trauma. It changes retinal structure into a wrinkle-like shape leading to visual impairment [2–4]. Clinically, ERM causes the loss of vision and metamorphopsia [5]. The prevalence of ERM is approximately 2% in people aged 60 years or below, and 12% in people aged 70 years and above [6]. Treatment for ERM includes removal of the ERM by vitreous surgery. Although improvements in visual acuity are attained in more than 75% patients who undergo vitreous surgery, metamorphopsia persists [3, 7–9].

Optical coherence tomography (OCT) is a noninvasive, non-contact imaging technique for in-depth visualization of the various layers of the retina [10]. Currently, advancements in OCT imaging have facilitated the capture of three-dimensional (3D) retinal structures within a few seconds with an axial resolution of $\sim$ 2 microns [11]. This 3D-OCT imaging can be performed by covering 360° of the retinal surface, and the 3D structure of the retinal surface layer can be stereoscopically displayed quantitatively by reconstructing multiple tomograms. Especially in the evaluation of ERM, surface maps of 3D-OCT can be useful because of the revealing manner and directions of membrane traction [12].

In recent years, image-processing technology using deep learning (DL) and support vector machine (SVM), machine-learning technologies, has attracted attention for its extremely high classification performance, and several studies for its applications in medical imaging have been reported [13–18]. Regarding the ophthalmic field, reports show that DL could be utilized to distinguish normal fundus images from retinal disease images using color fundus and OCT images [19, 20].

To the best of our knowledge, there have been no reports of the automatic diagnosis of ERM nor of machine-learning technologies using 3D-OCT images.

The aim of this study was to combine artificial intelligence with OCT for early disease detection. In this study, we performed a comparative review of SVM and DL, which are machine-learning technologies for detecting epithelial membrane using 3D-OCT images.

Methods

Data set

Of the patients diagnosed with ERM by two retinal specialists, those who were tested using 3D-OCT 2000 (Topcon Tokyo, Japan) were included. In addition, patients without fundus diseases were extracted from the clinical database of the Ophthalmology Department of Tsukazaki Hospital. These images were reviewed by two ophthalmologists for confirming the presence of ERM and were registered in an analytical database. In total, 529 3D-OCT images (184 non-ERM subjects and 205 ERM patients) were assessed; 80% of the images were divided for training and 20% for test as follows: 423 training (non-ERM 245, ERM 178) and 106 test (non-ERM 59, ERM 47) images. Using the 423 training images, a model was created using deep convolutional neural network (DNN) and SVM and the test data were evaluated. The present study complied with the Declaration of Helsinki. Research protocols and implementation were approved by the Ethics Committee of Tsukazaki Hospital.

3D-OCT imaging

To obtain images of the retinal surface layer expressing the height in the z-axis section, 3D-OCT was used to complement the images obtained by radial scan with 12 lines using visualization toolkit. Regarding the horizontal axis and area, the ratio adjustment was not special to unify the display (Fig. 1).

In the DNN analysis, right reversal treatment was applied to the left eye to ensure coincidence of the ear and nose sides.

3DL model

We implemented a DL model that uses a DNN in this classification system (see Fig. 2). For this, each convolutional layer was arranged in three layers. The activation function rectified linear unit (ReLU) and the max pooling layer (MP 1, 2, and 3) were placed after each convolutional layer [21]. Lastly, when the last two layers (FC 1, 2) were fully connected, we removed spatial information from extracted feature quantities and statistically recognized the target from other feature vectors. The last layer was a classification layer, using feature vectors of target images acquired in previous layers and the softmax function for binary classification. This network architecture is ideal for learning and recognition of the local features of complex image data with individual differences.

Training the DNN

All 3D-OCT images were converted to 96×96 pixels. For training, 100 images were mini-batch processed. The initial value of the network weight was randomly provided as the Gaussian distribution with a zero mean and a standard deviation of 0.05. To avoid overfitting during the training, a dropout technique

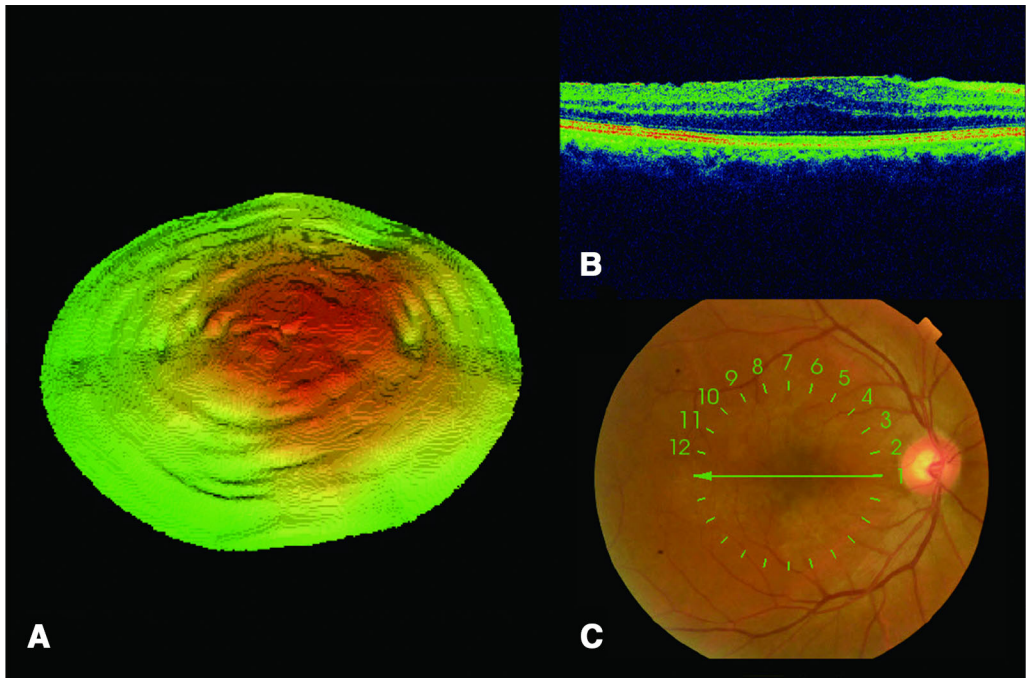

Fig. 1 **a** Representative 3D-OCT image of ERM generated. **b** OCT image of ERM, and **c** simultaneous scanning position and direction of the fundus photograph

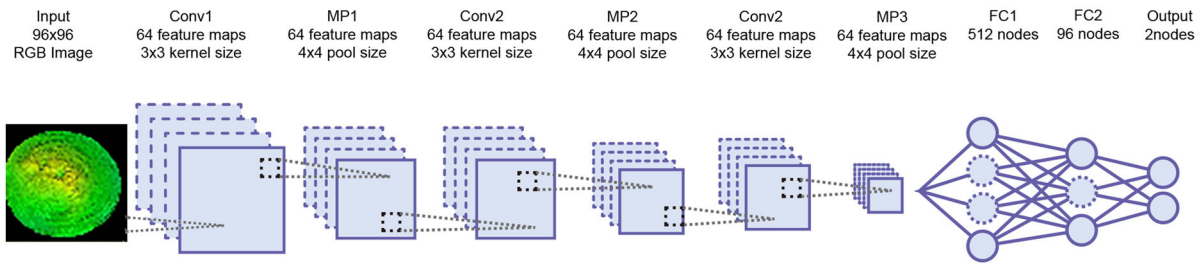

Fig. 2 Overall architecture of the model. The 96 × 96 pixel retinal fundus image data were described as input. Each of the convolutional layers (Conv1–3) is followed by an activation function (ReLU) layer, pooling layers (MP1–3) and two fully connected layers (FC1 and FC2). The final output layer performs binary classification using a softmax function

was applied to the fully connected layer 1 to mask out with 70% probability [22]. For improving the generalization performance, we updated the network weight according to the optimization algorithm [14–16]. Of the 100 DL models obtained in 100 learning cycles, those with the highest correct answer rate for the test data were selected as the DL model [22–24].

SVM

We used soft margin SVM, which is implemented in the scikit-learn library [25], with the radial basis function (RBF) kernel. The cost parameter (C) of SVM and the gamma parameter (γ) of RBF were grid searched for in the training data using quadrant cross-validation, and the combination with the highest average correct answer rate was selected. The parameter candidates were set as C = 1, 10, 100, and 1000, and γ = 0.0001, 0.001, 0.01, 0.1, and 1. We selected 10

dimensions with the highest correct answer rate for the test data by inputting the dimension reduction number as 2, 5, 10, 15, and 20. As a result, a learning model was generated using the parameter values of $C = 1$ and $\gamma = 0.001$.

Outcome

A receiver-operating characteristic (ROC) curve was created based on the discriminating ability of the DL and SVM model for normal and ERM 3D-OCT images and was evaluated using area under the curve (AUC), sensitivity, and specificity.

Statistical analysis

The 95% confidence interval (CI) of the ROC curve and AUC was obtained by the following method.

The images judged to exceed the cut-off value were defined as positive for ERM, and a ROC curve was created. The model was fitted to only 90% of the test data. Next, 100 ROC curves were generated by making 100 patterns, whereby 10% were thinned out, and then 100 AUCs were calculated from each ROC curve. Regarding AUC, a 95% CI was obtained by assuming normal distribution and using average/standard deviation. For sensitivity and specificity, optimal cut-off values calculated using the Youden index [26] were used as representative values in the DL model, using the first of the 100 ROC curves created, and the CI was calculated assuming a binomial distribution.

Results

Patient demographics

A total of 304 non-ERM images from 184 patients (mean age, 70.3 ± 8.7 years; 77 men and 107 women) and 225 ERM images from 205 patients (mean age, 70.1 ± 8.2 years; 79 men and 126 women) were analyzed (Table 1).

ROC curve and AUC analysis

The sensitivity and specificity of the DL model were 97.6% (95% CI, 87.7%–99.9%) and 98.0% (95% CI, 89.7%–99.9%), and the AUC was 0.993 (95% CI, 0.993–0.994). The sensitivity and specificity of the SVM model were 97.6% (95% CI, 87.7%–99.9%) and 94.2% (95% CI, 84.0%–98.7%), and the AUC was 0.988 (95% CI, 0.987–0.988) (Fig. 3).

DL and SVM model fitting

Using the 106 test data images, the DL model made two mistakes (1.9%; 1 non-ERM and 1 ERM), and the SVM model made five mistakes (4.4%; 3 non-ERM and 2 ERM). The mistakes made by the SVM model included all of the DL's mistakes. Figure 4 shows two images in which the SVM model made mistakes.

Discussion

Our results showed that the DL model for detecting ERM using the 3D-OCT images had higher AUC than the SVM model. DL is known to automatically learn local feature quantities of images and generate classification models [22, 24, 26, 27]. In addition, DL comprises many layers that are considered to learn and identify the local features of complicated individual differences, which can be combined [22]. A previous report compared spectral-domain OCT images of diabetic macular edema with normal eyes using SVM [28]. Recently, due to the evolution of DL, there are successive reports on the comparative advantages over SVM [29]. We and Asaoka et al. [30, 31] reported that DL was dominant when comparing SVM with DL.

Table 1 Patient demographics

	ERM	Non-ERM	*P* value
N	225	304	
Age	70.1 ± 8.2	70.3 ± 8.7	0.8638 (student's *t* test)
Sex (female)	126 (61.4%)	107 (58.1%)	0.535 (fisher's exact test)
Eye (left)	113 (49.8%)	157 (51.6%)	0.7257 (fisher's exact test)

ERM epiretinal membrane

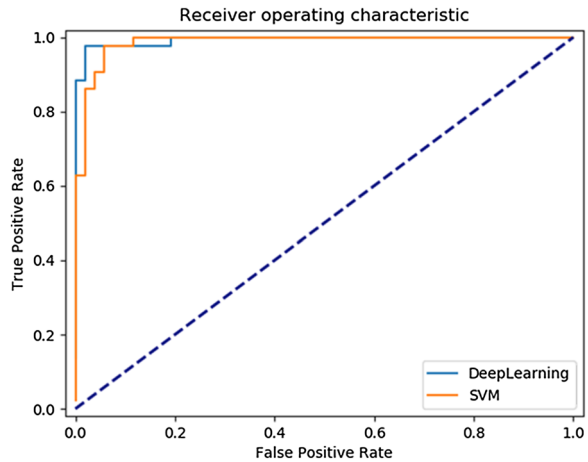

Fig. 3 Representative receiver-operating characteristic curves of the deep learning (DL) and support vector machine (SVM) models. The area under the curve (AUC) of the DL and SVM models were 0.993 (95% CI, 0.993–0.994) and 0.988 (95% CI, 0.987–0.988). The AUC was better for the DL model than for the SVM model

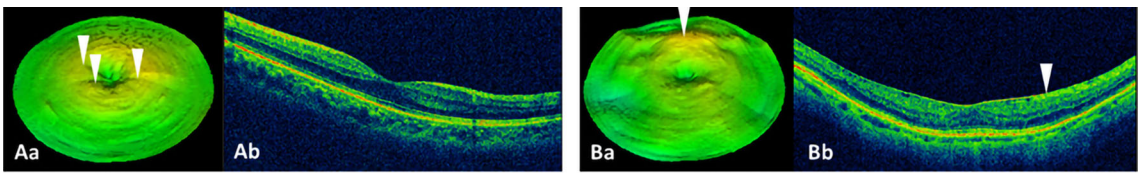

Fig. 4 Representative images, the SVM model was incorrect despite the correct answer by DL. *Aa* A 3D-OCT image from an non-ERM case, and *Ab* the OCT image of *Aa*. The SVM model identified this image as ERM. *Ba* A 3D-OCT image from an ERM case, and *Bb* the OCT image of *Ba*. The SVM model identified this image as non-ERM. In each case, the tomographic image was a reference and neither SVM nor DL was used as the identification information. The white arrowhead of *Aa* is a segmentation error, and discontinuity occurred due to the loss of the original thickness portion. Relatively clear retinal frontal membrane formation is observed (white arrow head in *Ba* and *Bb*)

In this study, the SVM model also showed a very high discrimination ability between ERM and non-ERM, but the DL model showed a higher numerical value than the SVM model. In test data images, in which SVM correctly answered, there was no image in which DL made a mistake. On the other hand, there was a case with relatively clear ERM that the DL model correctly answered, but the SVM model answered non-ERM. We believe this case indicated that the SVM model has a relatively low discrimination ability compared with the DL model and that it shows a high discrimination ability of the DL model.

For this study, we used 3D images instead of typical tomographic images on OCT because ERM lesions do not always spread 360° on the retinal surface layer. In fact, there is a possibility that ERM will not be detected on typical tomographic images with only the vertical and horizontal sections. However, with machine-learning technology, it is predicted that the analysis efficiency is significantly lowered when tomographic images were used for analysis in all 360° B-scans instead of typical sections because ERM exists on the upper layer of the retina. Therefore, 3D-OCT, which only represents the relative height information for 360° of the retinal surface layer, was selected. An additional advantage of 3D-OCT is that tomographic images are expressed while retaining a considerable part of curvature of the retina and individual differences of the horizontal axis; the horizontal axis and display range are standardized for 3D-OCT. Therefore, 3D-OCT is an optimal information source for qualitative identification of disease status. In this study, it was suggested that the 3D-OCT image has a great ability to discriminate between ERM and normal eyes regardless of whether the DL or SVM model was used.

The aim of this study was to combine artificial intelligence with OCT for early disease detection. If ERM were detected at an early stage, there is little adhesion and surgical treatment is relatively easy [32]. Therefore, there is a medical rationality for ERM screening during medical examinations using 3D-

OCT, which can be noninvasively performed under non-mydriatic exposure. However, it is economically inefficient to conduct 3D-OCT image interpretation for all acquired images by ophthalmologists. Therefore, the method of applying the DL model, as presented in this study, to the 3D-OCT interpretation for triage purposes may reduce the cost of such examinations, find at an early stage and result in the identification of more surgically adapted patients. Indeed, ERM surgery is performed based on not only OCT images but also physical symptoms such as loss of vision and metamorphopsia. Therefore, using DL model for triage and doing some tests (VA, *M*-charts which is a score of metamorphopsia) to those people can evaluate need for ERM surgery.

This study has certain limitations. The aim of this study was to examine the difference between the discriminating ability of DL models when applied to normal and ERM images on 3D-OCT; however, we could not show that the DL model can be applied in clinical settings. In addition, when clarity of the eye is reduced due to severe cataract or dense vitreous hemorrhage, it becomes challenging to captures images using 3D-OCT.

Conclusion

Both the DL and the SVM models exhibited sufficiently high capabilities to distinguish between ERM and normal eyes on 3D-OCT images. In addition, this study strongly suggested the possibility that the DL model is superior to the SVM model.

Acknowledgements I would like to thank M. Miki for creating the figures.

Compliance with ethical standards

Conflict of interest All authors certify that hey have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements) or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (Tsukazaki hospital) and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Okamoto F, Sugiura Y, Okamoto Y, Hiraoka T, Oshika T (2015) Inner nuclear layer thickness as a prognostic factor for metamorphopsia after epiretinal membrane surgery. Retina 35(10):1–8
2. Okamoto F, Okamoto Y, Hiraoka T, Oshika T (2009) Effect of vitrectomy for epiretinal membrane on visual function and vision-related quality of life. Am J Ophthalmol 147(5):869–874
3. Gomes LN, Corcostegui I, Fine HF, Chang S (2009) Subfoveal pigment changes in patients with longstanding epiretinal membranes. Am J Ophthalmol 147(5):865–868
4. Smiddy WE, Maguire AM, Green WR, Michels RG, de la Cruz Z, Enger C, Jaeger M, Rice TA (1989) Idiopathic epiretinal membranes. Ultra-structural characteristics and clinicopathologic correlation. Ophthalmology 96(6):811–821
5. Inoue M, Kadonosono K (2014) Macular diseases: epiretinal membrane. Dev Ophthalmol 54:159–163
6. Mitchell P, Smith W, Chey T, Wang JJ, Chang A (1997) Prevalence and association of epiretinal membranes. The Blue Mountains eye study, Australia. Ophthalmology 104(6):1033–1040
7. Bouwens MD, Meurs JC (2003) Sine Amsler Charts: a new method for the follow-up of metamorphopsia in patients undergoing macular pucker surgery. Grafes Arch Clin Exp Ophthalmol 241(2):89–93
8. Kinoshita T, Imaizumi H, Okushiba U, Miyamoto H, Ogino T, Mitamura Y (2012) Time course of changes in metamorphopsia, visual acuity, and OCT parameters after successful epiretinal membrane surgery. Invest Ophthalmol Vis Sci 53(7):3592–3597
9. Kim JH, Kang SW, Kong MG, Ha HS (2013) Assessment of retinal layers and visual rehabilitation after epiretinal membrane removal. Graefes Arch Clin Exp Ophthalmol 251(4):1055–1064
10. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA, Fujimoto JG (1991) Optical coherence tomography. Science 254(5035):1178–1181
11. Drexler W, Liu M, Kumar A, Kamali T, Unterhuber A, Leitgeb RA (2014) Optical coherence tomography today: speed, contrast, and multimodality. J Biomed Opt 19(7):071412
12. Legarreta JE, Gregori G, Knighton RW, Punjabi OS, Lalwani GA, Puliafito CA (2008) Three-dimensional spectral-domain optical coherence tomography images of the retina in the presence of epiretinal membranes. Am J Ophthamol 145(6):1023–1030
13. Lingyun G, Mingquan Y, Changrong W (2017) Cancer classification based on support vector machine optimized by

particle swarm optimization and artificial bee colony. Molecules 22(12):pii: E2086
14. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444
15. Liu S, Liu S, Cai W, Che H, Pujol S, Kikinis R, Feng D, Fulham MJ (2015) Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. IEEE Trans Biomed Eng 62:1132–1140
16. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, de Hulsbergen-van KC, Bult P, van Ginneken B, van der Laak J (2016) Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep 6:26286
17. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316:2402–2410
18. Pinaya WH, Gadelha A, Doyle OM, Noto C, Zugman A, Cordeiro Q, Jackowski AP, Bressan RA, Sato JR (2016) Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. Sci Rep 6:38897
19. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Proceedings of NIPS, pp 1097–1105
20. Gargeya R, Leng T (2017) Automated identification of diabetic retinopathy using deep learning. Ophthalmology 124(7):962–969
21. Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp 315–323
22. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. JMLR 15(1):1929–1958
23. Qian N (1999) On the momentum term in gradient descent learning algorithms. Neural Netw 12(1):145–151
24. Nesterov Y (1983) A method for unconstrained convex minimization problem with the rate of convergence O (1/k^ 2). Doklady AN USSR 269:543–547
25. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830
26. Schisterman EF, Faraggi D, Reiser B, Hu J (2008) Youden Index and the optimal threshold for markers with mass at zero. Stat Med 27(2):297–315
27. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009) ImageNet: a large-scale hierarchical image database. In: IEEE conference on computer vision and pattern recognition, CVPR 2009
28. Alsaih K, Lemaitre G, Rastgoo M (2017) Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT image. BioMed Eng Online 16:68
29. Liu B, Liu Y, Zhou K Image classification for dogs and cats. https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf
30. Ohsugi H, Tabuchi H, Enno H, Ishitobi N (2017) Accuracy of deep learning, a machine-learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment. Sci Rep 7(1):9425. https://doi.org/10.1038/s41598-017-09891-x
31. Asaoka R, Murata H, Iwase A, Araie M (2016) Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. Ophthalmology 123(9):1974–1980
32. Kim JS, Chhablani J, Chan K, Cheng L, Kozak I, Hartmann K, Freeman WR (2012) Retinal adherence and fibrillary surface changes correlate with surgical difficulty of epiretinal membrane removal. Am J Ophthalmol 153(4):692–697