Int Ophthalmol (2019) 39:1269–1275
https://doi.org/10.1007/s10792-018-0940-0

ORIGINAL PAPER

Accuracy of ultra-wide-field fundus ophthalmoscopy-assisted deep learning, a machine-learning technology, for detecting age-related macular degeneration

Shinji Matsuba · Hitoshi Tabuchi · Hideharu Ohsugi · Hiroki Enno · Naofumi Ishitobi · Hiroki Masumoto · Yoshiaki Kiuchi

Received: 29 January 2018 / Accepted: 2 May 2018 / Published online: 9 May 2018

Abstract
Purpose To predict exudative age-related macular degeneration (AMD), we combined a deep convolutional neural network (DCNN), a machine-learning algorithm, with Optos, an ultra-wide-field fundus imaging system.
Methods First, to evaluate the diagnostic accuracy of DCNN, 364 photographic images (AMD: 137) were amplified and the area under the curve (AUC), sensitivity and specificity were examined. Furthermore, in order to compare the diagnostic abilities between DCNN and six ophthalmologists, we prepared yield 84 sheets comprising 50% of normal and wet-AMD data each, and calculated the correct answer rate, specificity, sensitivity, and response times.
Results DCNN exhibited 100% sensitivity and 97.31% specificity for wet-AMD images, with an average AUC of 99.76%. Moreover, comparing the diagnostic abilities of DCNN versus six ophthalmologists, the average accuracy of the DCNN was 100%. On the other hand, the accuracy of ophthalmologists, determined only by Optos images without a fundus examination, was 81.9%.
Conclusion A combination of DCNN with Optos images is not better than a medical examination; however, it can identify exudative AMD with a high level of accuracy. Our system is considered useful for screening and telemedicine.

Keywords Ultra-wide-field scanning laser ophthalmoscope · Neural networks · Age-related macular degeneration · Pattern recognition · Telemedicine

S. Matsuba (✉) · H. Tabuchi · H. Ohsugi · N. Ishitobi · H. Masumoto
Department of Ophthalmology, Saneikai Tsukazaki Hospital, 68-1 Aboshi Waku, Himeji 671-1227, Japan
e-mail: s.matsuba@tsukazaki-eye.net

S. Matsuba · Y. Kiuchi
Department of Ophthalmology and Visual Sciences, Graduate School of Biomedical Sciences, Hiroshima University, 1-2-3 Minami, Kasumi, Hioroshima 734-8553, Japan

H. Enno
Rist Inc., 2-11-3 Meguro, Meguro-ku, Tokyo 153-0063, Japan

Introduction

Age-related macular degeneration (AMD) has been reported as the third-most common causes of blindness among the elderly population in developed countries [1]. In Japan, the incidence of AMD has increased in recent years [2]. Although the introduction of anti-vascular endothelial growth factor (anti-VEGF) therapy for the treatment of AMD has facilitated the maintenance and improvement of visual acuity [3, 4], some cases involving prolonged exudation deteriorate to macular atrophy despite receiving anti-VEGF

agents [5, 6]. In addition, the long-term visual prognosis following anti-VEGF therapy depends on the patient's age and visual acuity at treatment initiation [7], thus underscoring the importance of an early stage specialized ophthalmic consultation and the appropriate timing of treatment. This timing is complicated by an insufficient number of clinical specialists in age-related macular degeneration worldwide. Even in Japan, only 20.6% of ophthalmologists belong to the professional society of the retina and more than 5 years of clinical experience are required before a clinician is approved to administer PDT treatment. In other words, professional skills training is a lengthy process.

Currently, the Optos 200 Tx (Optos PLC, Dunfermline, United Kingdom) is the most widely used commercially available ultra-wide-field scanning laser ophthalmoscope. As shown in Fig. 1, this noninvasive device can capture a wide range of fundus photographs over a viewing angle of 200°, even under non-mydriasis, and can cover 80% of the fundus in a single shot [8].

The Optos does not require a mydriatic agent and, thus, avoids risk of pupillary block as well as the requirement for an ophthalmologist's review before imaging. Several reports have discussed the role of ultra-wide-field scanning in the diagnosis, follow-up, and treatment predictions in the context of various fundus diseases [9–12].

Recent studies have reported progress in the use of image processing technologies based on deep learning algorithms, which have very high image discrimination abilities, for the analyses of medical images [13–17]. Regarding AMD, a previous report described the use of AREDS data to perform staging from fundus camera images and observed similar levels of precision when comparing the discriminatory abilities of a deep learning algorithm with those of physicians [18]. However, that report aimed to substitute medical treatment at the level of an AMD specialist and used image data set comprising the fundus photographs obtained under mydriasis and confined to the macula.

Our research group previously reported that a combination of the high discrimination ability of a deep convolutional neural network (DCNN), a deep learning algorithm, with Optos allowed us to detect early retinal detachment in the context of ophthalmic telemedicine [19]. In addition, this result will contribute to the reduction of social security cost which is a heavy burden in many countries worldwide [20]. We believe that we are the first group to report the indispensability of the Optos for telemedicine. Here, we examined the accuracy of a combination of DCNN and Optos images to distinguish between exudative AMD and a normal eye and compared the diagnostic ability of an ophthalmologist examination with that of our combined technique.

Methods

Dataset

This study was conducted in compliance with the Declaration of Helsinki. The research protocols and implementation were approved by the ethics committee of Tsukazaki Hospital.

We extracted 227 Optos images of normal patients without fundus disease (normal data) and 137 Optos images of patients with exudative AMD (wet-AMD data) from the clinical ophthalmology database at Tsukazaki Hospital. To prevent significant differences between the two groups of normal and wet-AMD, we intentionally controlled the age parameter in the

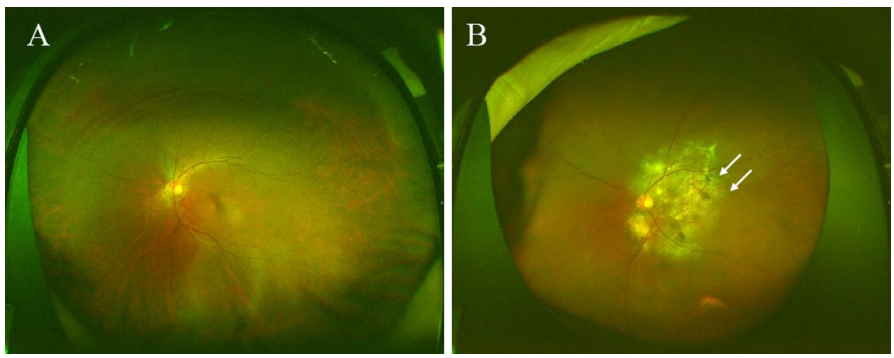

Fig. 1 Representative fundus image obtained via ultra-wide-field scanning laser ophthalmoscopy. Ultra-wide-field fundus images of a normal left eye (**a**), and left eye with AMD (**b**). Arrows indicate retinal atrophy

normal group; males were aged > 70 years and females > 77 years.

We divided each image data group such that 70 and 30% of images were included in training and test datasets, to yield 253 training images (normal: 158, wet-AMD: 95) and 111 test images (normal: 69, AMD: 42). Subsequently, the 253 training images were subjected to contrast adjustment, gamma correction, histogram equalization, noise addition, and image reversal processing, followed by amplification to 5000 images (normal: 4130 sheets, wet-AMD: 870 sheets). We further evaluated the wet-AMD data to diagnostic confirmation based on the results of ordinary fundus examinations, OCT, and FA/IA examinations conducted by two independent retinal specialists. As we aimed to determine the presence or absence of AMD, we excluded cases with unclear images attributed to vitreous hemorrhage, astrocytosis, or strong cataracts and cases with previous retinal photocoagulation and other complicating fundus diseases as determined by retinal specialists.

Deep learning model

We implemented a DCNN-based deep learning model as shown in Fig. 2. After arranging three convolutional layers, we arranged the activation functions of rectified linear unit (ReLU) [21] and batch normalization [22] is arranged. Max pooling layers (MP 1, 2) were placed after convolutional layers 1 and 3, and a dropout layers (drop rate: 0.25) were placed after each MP. Finally, two full connection layers (FC 1, 2) comprising all layers were arranged and separated into two classes using the softmax function.

Training the DCNN

All image data were initially converted to a pixel resolution of 256 × 192. Training involved the mini-batch processing of 10 images with an epoch number of 100 times. The initial value of the network weight was randomly given as the zero average of the Gaussian distribution, with a standard deviation of 0.05. Dropout processing was performed to mask the first total tie layer (FC1) with 50% probability [23]. This weight was updated according to the optimization algorithm [momentum SGD (learning coefficient = 0.01, inertia term = 0.9)] [24, 25]. Of the 100 deep learning models obtained during 100 learning cycles, the most correct rate that was higher in test data was selected as deep learning model. For this study, a system with Windows 10 Home, an Intel Core i7—3630 QM CPU, and 8.00 GB of memory was used.

Outcome

We examined the area under the curve (AUC), sensitivity, and specificity for the ability of DCNN to discriminate between normal and wet-AMD data as described above.

Statistical analysis

The receiver operating characteristic (ROC) curve used to determine the AUC was created by defining the point at which the value used to indicate wet-AMD positively exceeds the threshold (cutoff value) output from the softmax function. By creating 100 ROC curves from 100 patterns and thinning out 10%, this

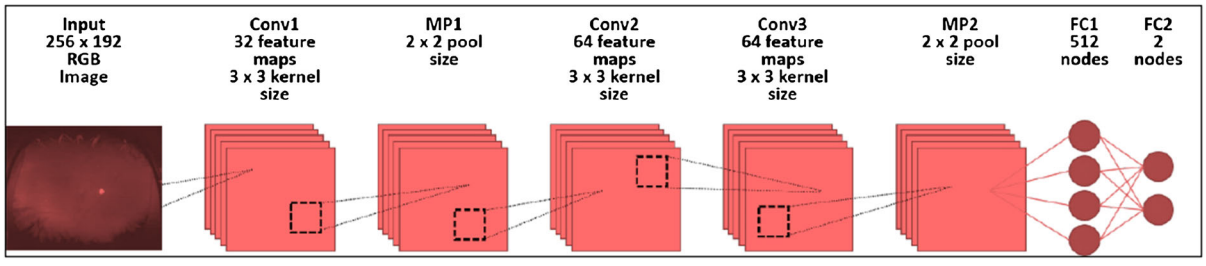

Fig. 2 Overall architecture of the model. The image data were converted to a pixel resolution of 256 × 192 and set as the input. After placing the convolution layers (Conv 1, 2, 3), activation function (ReLU), pooling layers (MP 1, 2) after Conv 1 and 3, and a dropout layer (drop rate: 0.25), all were passed through two full coupling layers (FC 1 and 2). In the final output layer, the classification was performed using two class softmax functions

model was applied to only 90% of the test data. One hundred AUCs were calculated from each ROC curve and a 95% confidence interval was obtained by assuming a normal distribution and average standard deviation. The sensitivity and specificity at the optimal cutoff value, calculated using the Youden index [26], were used as a representative value of the deep learning model, using the first one among 100 ROC curves.

The confidence intervals of sensitivity and specificity were calculated, assuming a binomial distribution. We used JMP Pro, version 10.0.0 (SAS Institute Inc., Cary, NC, USA) for the data analysis.

Creation of the test application for ophthalmologist interpretation

As the second experiment, we compared the diagnostic accuracy between DCNN and ophthalmologists. For this, we prepared 111 images (normal: 69, wet-AMD: 42) as the test data. Of these, 42 images were extracted from normal data using a random number generation method to yield 84 sheets comprising 50% of normal and wet-AMD data each. A total of 84 sheets were used as test data. The correct answer rate, specificity, sensitivity, and response times by DCNN and six ophthalmologists were calculated.

Determination and required time measurement methods

Six ophthalmologists determined the presence or absence of AMD looking at 84 test data images presented on a computer monitor, without a fundus examination. Each examine entered the integer 0 or 1 in a Microsoft Excel-based response form. To determine the average time required for the ophthalmologist to reply, the time to response entry was also included on the computer monitor.

The DCNN performed a series of tasks, including confirming the number of the problem in the answer column → reading the image → judging → filling all presented numbers in the answer columns, and counting the total time as the operation time. This series of work was performed 15 times by the computer and its median value was recorded as the working time.

Heatmap creation

It is illustrated as a heatmap, which DCNN focuses on which coordinate axes on the image are being classified. The heatmap was generated using Gradient-weighted Class Activation Mapping (Grad-CAM) [27] and a gradient layer using convolution layer 2 was designated. ReLU was specified as the backprop_modifier.

Result

A total of 227 normal data images from 156 patients (mean age: 77.0 ± 4.8 years; 113 men and 43 women) and 137 wet-AMD data images from 114 patients (mean age: 76.2 ± 8.2 years; 84 men and 30 women) were included. The two groups of patients did not differ significantly in terms of age, sex ratio, or left/right affected eye ratio ($P = 0.6252$, $P = 0.6247$, and $P = 0.9139$, respectively; Tukey–Kramer test and Fisher's exact test) (Table 1).

Performance of DCNN

DCNN yielded an average AUC of 99.76% (95% CI: 99.75–99.76) (Fig. 3), with an average sensitivity of 100% (95% CI: 90.97–n/a) and specificity of 97.31% (95% CI: 93.84–99.12).

Comparison of ophthalmologist and DCNN judgments of test images

In the analysis of 84 test images (normal data: 42, wet-AMD data: 42), the six ophthalmologists yielded an average correct answer rate of 81.9%, sensitivity of 71.4% and specificity of 92.5%. The average required time was 11 min, 23.54 s. The DCNN yielded an average correct answer rate, sensitivity, and specificity of 100%, with a required time of 0 min, 26.29 s (Table 2).

Heat map

A typical heat map is shown in Fig. 4. The areas of strong DCNN focus are indicated in pale blue color. The focal points accumulated throughout the lesion.

Table 1 Comparison of demographic variables between the normal and wet-AMD data groups

	Wet-AMD	Normal	*P* value	Test
Eyes	137	227		
Mean age, years (range)	76.2 ± 8.2 (48–92)	77.0 ± 4.8 (70–94)	0.6252	Tukey–Kramer
Sex, female	34 (24.8%)	63 (27.8%)	0.6247	Fisher's exact
Eyes, left	66 (48.18%)	112 (49.34)	0.9139	Fisher's exact

No statistically significant differences were observed between the groups. Data are presented as numbers (%) unless otherwise indicated

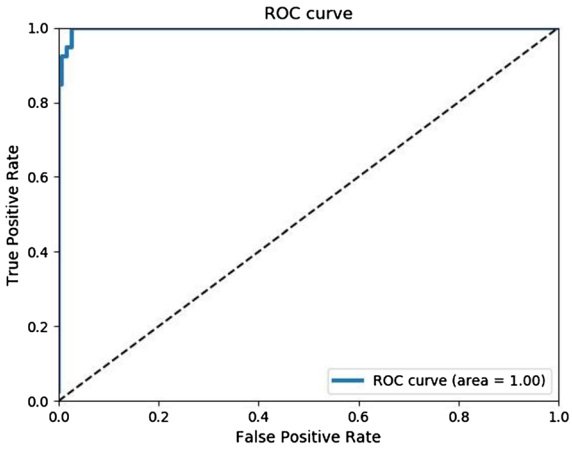

Fig. 3 The optimum receiver operating characteristic (ROC) curve yielded by DCNN. This curve was selected from among the 100 created curves

Table 2 Comparison of the abilities of the DCNN and ophthalmologists (n = 6) to distinguish normal from Wet-AMD data

	DCNN	Ophthalmologists
Correct answer rate	100%	81.9 ± 4.8%
Sensitivity	100%	71.4 ± 13.9%
Specificity	100%	92.5 ± 6.5%
Required time	0:26.29	11:23.54 ± 2:42.20

Discussion

In this study, a combination of DCNN and Optos image data could distinguish between normal and wet-AMD data with a high level of accuracy. Furthermore, in a comparison between the abilities of DCNN and ophthalmologists to distinguish between normal and wet-AMD data, the former outperformed the latter in terms of the correct answer rate, specificity, sensitivity, and response speed.

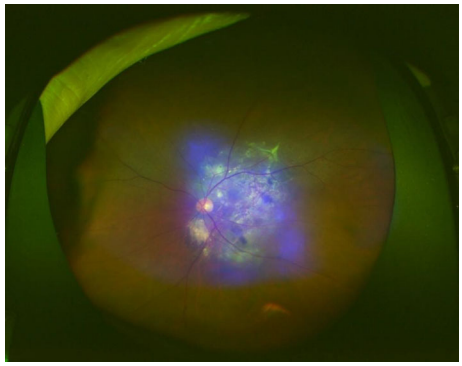

Fig. 4 An image of wet-AMD from Fig. 1, overlaid with the created heatmap. The areas of DCNN focus are indicated in pale blue. The DCNN clearly evaluated the entire AMD lesion

The multilayered DCNN automatically learns the local features of an image and generates the classification model. Accordingly, this network can conceive and construct an optimum structure for learning and identifying the local features of complex and individual differences in image data [28–30]. The convolutional layer acquires the amounts of features of an object using a convolution filter. Although MP layers are placed after convolution layers 1 and 3, these aim to reduce the position sensitivity of the feature amount outputted from the convolution layer and facilitate more general recognition. The dropout layer installed after each MP layer aims to avoid excess adaptation (i.e., overlearning) to the training data. Finally, the FC layers 1 and 2 remove spatial information from the extracted feature amount and statistically identify the object from other feature vectors. Dropout processing to mask the first total tie layer (FC1) with 50% probability was performed to improve generalizability and ensure that overlearning would not occur [23].

Since the introduction of anti-VEGF agents, meta-analyses have shown that treatment can maintain visual function in patients with AMD over the long term [31]. As noted previously, an early stage diagnosis and appropriately timed treatment can help to maintain better vision, as long-term visual prognosis worsens with increasing patient age and worse visual acuity at treatment initiation [7]. Therefore, early diagnosis is very important for maintaining good visual function in patients with AMD. As noted above, the Optos can safely, easily, and non-invasively capture wide-area fundus images in the absence of an ophthalmologist [20] and could be used to identify exudative AMD, as proposed in this study in combination with DCNN. The use of such a system at a medical center lacking an ophthalmologist would likely increase probability of detecting AMD without subjective symptoms. In addition, the use of this technology in remote locations (i.e., far from urban areas) lacking ophthalmological services, the high identification capability of this technology may encourage subsequent visits to ophthalmologists. As blindness prevention has a socially significant economic effect [32], the placement of such a system in an examination center or other area lacking an ophthalmologist could improve the early detection of AMD and reduce the incidence of blindness.

However, this study had several limitations. First, it is difficult to acquire precise images using Optos when the transmission of light into the eye is impaired by an intermediate translucent zone (e.g., intense cataract, vitreous cloudiness). Accordingly, patients with this characteristic were excluded from our study. Second, we excluded patients with fundus hemorrhagic diseases such as diabetic retinopathy and retinal vascular occlusive diseases, which require careful discrimination from AMD, as well as patients with complications of glaucoma. Furthermore, the wet-AMD dataset comprised all relatively active cases of exudative AMD, most of which had been treated using anti-VEGF vitreous injection therapy. Future studies should involve larger numbers of cases and should aim to evaluate early and slow progressive AMD using the same procedure described in this study. The results would clarify the versatility and effectiveness of this method.

Finally, six ophthalmologists have determined the presence or absence of AMD, without a fundus examination. Therefore, although combination of DCNN and Optos images is certainly better, it is not particularly superior to a medical examination. It should be noted that the inspection of a face-to-face by ophthalmologists is indispensable for a definite diagnosis. In addition, both conventional OCT and angiography, by retinal specialists, are essential to confirm a qualitative diagnosis, treatment effects, and follow-up observations.

Conclusion

In this study, we revealed that a combination of DCNN and Optos had a high discriminating ability to detect exudative AMD. The Optus is a non-invasive tool, and our system, which combines DCNN with Optos, is useful for the purposes of screening and telemedicine. In addition, this result may contribute to the reduction of social security cost.

Acknowledgements The authors thank Masayuki Miki and orthoptists of Tsukazaki Hospital for support in data collection.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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 and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

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

References

1. Lim LS, Mitchell P, Seddon JM, Holz FG, Wong TY (2012) Age-related macular degeneration. Lancet 379:1728–1738
2. Yasuda M, Kiyohara Y, Hata Y, Arakawa S, Yonemoto K, Doi Y, Iida M, Ishibashi T (2009) Nine-year incidence and risk factors for age-related macular degeneration in a defined Japanese population the Hisayama study. Ophthalmology 116:2135–2140
3. Rosenfeld PJ, Brown DM, Heier JS, Boyer DS, Kaiser PK, Chung CY, Kim RY, MARINA Study Group (2006) Ranibizumab for neovascular age-related macular degeneration. N Engl J Med 355:1419–1431
4. Brown DM, Michels M, Kaiser PK, Heier JS, Sy JP, Ianchulev T, ANCHOR Study Group (2009) Ranibizumab versus verteporfin photodynamic therapy for neovascular

age-related macular degeneration: two-year results of the ANCHOR study. Ophthalmology 116:57–65
5. Comparison of Age-related Macular Degeneration Treatments Trials (CATT) Research Group, Maguire MG, Martin DF, Ying GS, Jaffe GJ, Daniel E, Grunwald JE, Toth CA, Ferris FL, Fine SL (2016) Five-year outcomes with anti-vascular endothelial growth factor treatment of neovascular age-related macular degeneration. Ophthalmology 123(8):1751–1761
6. Rofagha S, Bhisitkul RB, Boyer DS, Sadda SR, Zhang K, SEVEN-UP Study Group (2013) Seven-year outcomes in ranibizumab-treated patients in ANCHOR, MARINA, and HORIZON: a multicenter cohort study (SEVEN-UP). Ophthalmology 120(11):2292–2299
7. Rasmussen A, Bloch SB, Fuchs J, Hansen LH, Larsen M, Lacour M, Lund-Andersen H, Sander B (2013) A 4-year longitudinal study of 555 patients treated with ranibizumab for neovascular age-related macular degeneration. Ophthalmology 120:2630–2636
8. Witmer MT, Parlitsis G, Patel S, Kiss S (2013) Comparison of ultra-widefield fluorescein angiography with the Heidelberg Sp ectralis(®) noncontact ultra-widefield module versus the Optos(®) Optomap(®). Clin Ophthalmol 7:389–394
9. Nagiel A, Lalane RA, Sadda SR, Schwartz SD (2016) Ultra-widefield fundus imaging: a review of clinical applications and future trends. Retina 36:660–678
10. Ghasemi Falavarjani K, Tsui I, Sadda SR (2017) Ultra-wide-field imaging in diabetic retinopathy. Vis Res S0042–6989(17):30118–30119
11. Soliman AZ, Silva PS, Aiello LP, Sun JK (2012) Ultra-wide field retinal imaging in detection, classification, and management of diabetic retinopathy. Semin Ophthalmol 27(5–6):221–227
12. Witmer MT, Cho M, Favarone G, Chan RV, D'Amico DJ, Kiss S (2012) Ultra-wide-field autofluorescence imaging in non-traumatic rhegmatogenous retinal detachment. Eye (Lond) 26(9):1209–1216
13. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444
14. Liu S, Liu S, Cai W, Che H, Pujol S, Kikinis R, Feng D, Fulham MJ, ADNI (2015) Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. IEEE Trans Biomed Eng 62:1132–1140
15. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-van de Kaa C, 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
16. 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
17. 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
18. Burlina P, Pacheco KD, Joshi N, Freund DE, Bressler NM (2017) Comparing humans and deep learning performance for grading AMD: a study in using universal deep features and transfer learning for automated AMD analysis. Comput Biol Med 82:80–86
19. Mrsnik M (2013) Global aging 2013: rising to the challenge. Standard and poor's rating services network. https://www.nact.org/resources/2013_NACT_Global_Aging.pdf
20. 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 7:9425
21. Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: The 14th international conference on artificial intelligence and statistics, pp 315–323
22. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning
23. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15:1929–1958
24. Qian N (1999) On the momentum term in gradient descent learning algorithms. Neural Netw 12(1):145–151
25. Nesterov Y (1983) A method for unconstrained convex minimization problem with the rate of convergence O (1/k^2). Doklady AN USSR. 269
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. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2016) Grad-CAM: visual explanations from deep networks via gradient-based localization. https://arxiv.org/abs/1610.02391v3
28. Deng J, Dong W, Socher R, Li LJ, Li K, Li FF (2009) Imagenet: a large-scale hierarchical image database. In: Computer vision and pattern recognition, pp 248–255
29. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Li FF (2014) Imagenet large scale visual recognition challenge. arXiv preprint arXiv:1409.0575
30. Lee CY, Xie S, Gallagher P, Zhang Z, Tu Z (2016) Deeply-supervised nets. arXiv preprint arXiv:1409.5185 (2014)
31. Gerding H (2016) Long-term results of intravitreal Anti-VEGF injections in wet AMD: a meta-analysis. Klin Monbl Augenheilkd 233:471–474
32. Gordois A, Cutler H, Pezzullo L, Gordon K, Cruess A, Winyard S, Hamilton W, Chua K (2012) An estimation of the worldwide economic and health burden of visual impairment. Glob Public Health 7(5):465–481