Intelligence-Based Medicine 3-4 (2020) 100024

Contents lists available at ScienceDirect

Intelligence-Based Medicine

journal homepage: www.journals.elsevier.com/intelligence-based-medicine

Automatic screening for diabetic retinopathy in interracial fundus images using artificial intelligence

Yusaku Katada [a,b,1], Nobuhiro Ozawa [a,b,1], Kanato Masayoshi [a], Yoshiko Ofuji [b], Kazuo Tsubota [b,**], Toshihide Kurihara [a,b,*]

[a] Laboratory of Photobiology, Keio University School of Medicine, Shinanomachi, Shinjuku-ku, Japan
[b] Department of Ophthalmology, Keio University School of Medicine, Shinanomachi, Shinjuku-ku, Japan

A R T I C L E I N F O

Keywords:
Artificial intelligence
Diabetic retinopathy
Fundus photography

A B S T R A C T

The development of deep learning technology and high-performance computers in recent years has enabled significant improvement in the accuracy of image recognition, and efforts are currently being made toward the development of ophthalmic artificial intelligence (AI). In this study, deep learning was applied to fundus image analysis. The fundus images were automatically classified by their severity index and the recognition accuracies for Japanese and American data were compared. We trained a deep convolutional neural network and a support vector machine using a data set consisting of 35,126 open-source American clinical fundus images, which were graded for diabetic retinopathy by licensed ophthalmologists. The grading accuracy of the trained AI model was evaluated by comparison with its performance for 200 Japanese fundus images obtained at Keio University Hospital. The trained AI model exhibited a sensitivity of 81.5% and specificity of 71.9% for the American validation data set, and a sensitivity of 90.8% and specificity of 80.0% for the Japanese data set. This indicates that the proposed AI program developed using American image data can be applied to fundus photographs of Japanese subjects and thus represents an interracial screening model. It can be used for screening tests and applied to telemedicine systems.

1. Introduction

Diabetic retinopathy (DR) is a characteristic complication of diabetes that causes vision impairment, accounting for 19.0% (third causative factor) of visual impairment in Japan [1]. DR results from long-standing diabetes and is conventionally clinically identified by the presence of lesions associated with vascular abnormalities through direct fundus examination or fundus photographs. Early and routine medical checkups are very important for dealing with DR because the disease often shows few symptoms until it is too late for effective treatment. However, there is an inadequacy of experienced ophthalmologists that can perform direct periodic medical checkup for fundus and evaluate the fundus photographs. Furthermore, the examination process is time-consuming and there is a greater scarcity of experienced physicians in local areas where there are more diabetic patients. There is thus the need for an effective infrastructure for preventing blindness in the increasing number of diabetic patients globally.

The development of deep learning technology and high-performance computers in recent years has enabled significant improvement in the accuracy of image recognition. By using a large amount of image data and correct labeling, it is possible to classify images without explicitly defining their features quantitatively. This technology has the potential for use in developing a highly reliable artificial intelligence (AI) system for reducing the human cost of the diagnosis and severity evaluation of DR based on fundus photographs. Such an AI system promises to improve ophthalmic public health delivery in areas where ophthalmologists are less available, and relevant research efforts are already ongoing. Recent study findings suggest that excellent diagnostic and evaluation accuracy can be achieved by the procedure [2–7] using the same dataset. However, the performance and robustness of the system for other racial populations such as Japanese are currently unknown. This is because the accuracy of some of the tasks such as the automatic calculation of the

* Corresponding author. Laboratory of Photobiology, Keio University School of Medicine, Shinanomachi, Shinjuku-ku, Japan.
** Corresponding author.
E-mail addresses: yusakukatada@keio.jp (Y. Katada), n.ozw@z5.keio.jp (N. Ozawa), kanato_masayoshi@keio.jp (K. Masayoshi), ofuji.yoshiko413@keio.jp (Y. Ofuji), tsubota@z3.keio.jp (K. Tsubota), kurihara@z8.keio.jp (T. Kurihara).

[1] These authors have contributed equally.

https://doi.org/10.1016/j.ibmed.2020.100024
Received 13 July 2020; Received in revised form 15 November 2020; Accepted 3 December 2020
2666-5212/© 2020 Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

retinal vascular diameter may vary across races. Specifically, it is not clear whether the classification task of an AI system trained using data obtained from a particular race would be accurate for another races.

In the present study, we evaluated the fundus images of Japanese subjects using an AI image classification model developed using open-source American fundus images with the purpose of comparing the interracial errors.

2. Material and methods

2.1. Data set

The model algorithm was developed using different EyePACS data obtained from Kaggle (https://www.kaggle.com/c/diabetic-retinopathy-detection/data), which is a collection of 35,126 images of DR collected in the United States and labelled into five classes. Because multiple images were acquired from the same patients, the utilized images were randomly selected when the classifications were the same. The fundus images were acquired by various cameras (https://www.eyepacs.com/solution#cameras) with field views of 40°–50°. The validation data set for Japanese subjects consisted of 200 fundus photographs obtained from 200 patients who visited the Keio University Hospital, Tokyo, Japan for the first time between January 2012 and May 2015. The photographs were acquired by a TRC-50DX retinal camera (Topcon, Tokyo, Japan) with a field view of 50°. The Keio dataset consisted of all Japanese patients with an average age of 67.9 (±12.6) and 23.5% female. The demographic variables of this public EyePACS dataset are unknown. For reference, another EyePACS dataset averaged 54.1 (±11.3) years of age, with 62.2% of female and was enriched for Hispanic patients (~55%), with Caucasian, Black, and Asian patients each comprising approximately 5–10% of the population [3].

The study was approved by the institutional review board of the Keio University School of Medicine (Approval No. 20170049). All the training and clinical validation images were graded for DR by three licensed Japanese ophthalmologists. The DR severity (none, mild, moderate, severe, or proliferative) was graded according to the International Clinical Diabetic Retinopathy scale [8]. A moderate or worse DR was defined as a referable DR (Michael D [9]. All the images were graded based on one posterior fundus image with a 40°–50° field of view. Those with varying grading results were repeatedly graded until they finally matched until constant results were obtained. Images of poor quality were eliminated.

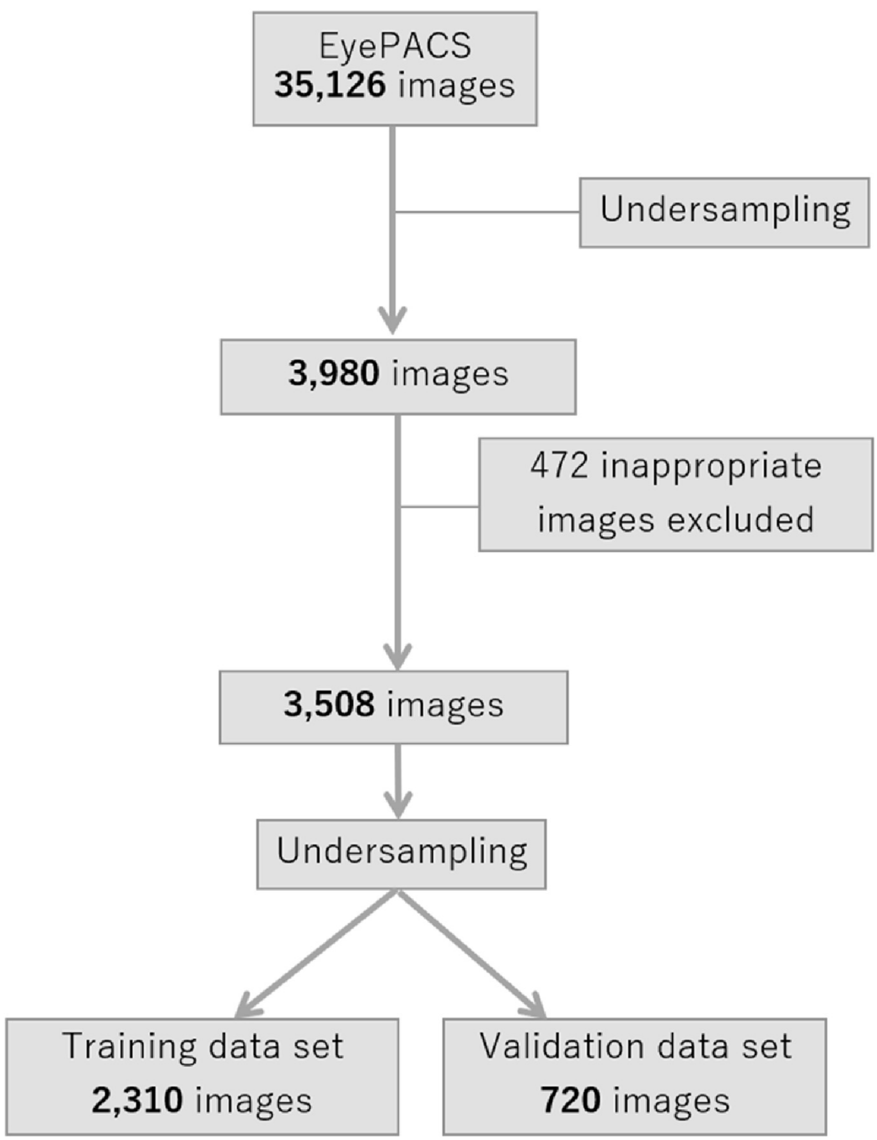

Fig. 1. Arrangement of the EyePACS data for detection of diabetic retinopathy.

2.2. Data selection

Due to class imbalance and the presence of some low-quality images in the original EyePACS dataset, the utilized images were selected as follows (Fig. 1). First, 3980 images were randomly selected for data curation. Second, based on the data curation, 472 inappropriate images were excluded, leaving 3508 images. Finally, from each DR severity class, 462 images were randomly selected as the EyePACS training dataset, and 144 as EyePACS validation data. Thus, 2310 images were used for the training of the model, and 720 images for its evaluation. The Keio image data was collected in chronological order and included 40 samples of each grade. Images that were blurred due to cataracts and vitreous opacity had been excluded. Because the purpose of this study was to examine the adaptability of the AI model developed using American fundus images to Japanese subjects, the number of samples of each grade was under sampled.

2.3. Preprocessing

The field of view of each image was first extracted. In the training step, data augmentation was conducted including horizontal flipping, random brightening, random contrasting, random saturation, random rotation (by $\leq 2^{\circ}$), and resizing to 512×512 pixels (using the three RGB channels).

2.4. Training algorithm

A support vector machine (SVM) and neural network (NN) were used for the binary classification task (determination of whether the DR was referable or not) and the multiclass classification task (into one of the five DR severity classes). The SVM classifier had an Inception v3 CNN architecture (Google, CA, United States) that was pretrained on the ImageNet dataset [10,11] for extraction of the image features. The final classification layer was subsequently removed from the network and a global average pooling (GAP) layer, which calculates the average of each feature map of the final convolution layer, was added as a replacement to reduce the number of features and facilitate the learning. The weights of the Inception v3 layers were fixed. The 2048 extracted features were used to train the support vector machine. The binary classifier was used to conduct Platt scaling [12] to determine the classification probabilities. An image with an output probability higher than 0.5 was classified as referable.

The neural network (NN) classifier had an Inception v3 CNN architecture as well as a GAP layer, fully connected layers, and either a softmax function (for a multiple classification model) or sigmoid function (for a binary classification model). Unlike the SVM classifier, the weights of the Inception v3 layers and fully connected layers were trainable, but the initial weights of the Inception v3 layers were set to the pretrained weights.

2.5. Evaluation

We trained the SVM (binary classification), NN (binary classification), SVM (multiple classification), and NN (multiple classification) using the EyePACS training dataset. The performances of each model for the EyePACS validation dataset and Keio dataset were compared to examine the model's adaptability. The evaluation metrics were the confusion matrices, sensitivity, specificity, and ROC-AUC, which is the area under the receiver operating characteristic (ROC) curve (AUC denotes area under curve). In addition, to further interpret the NN model, we showed the gradient-weighted class activation mapping (Grad-CAM) [13] of some samples in EyePACS and Keio dataset.

3. Results

The statistical details of the images are summarized in Table 1.

The performances of the SVM classifiers for DR detection in the EyePACS and Japanese (Keio) validation data are summarized in Figs. 2 and 3 and Table 2. The total accuracy of five-class SVM classifier was determined to be 48.8 ($\pm$3.7) % for the EyePACS validation data and 40.0 ($\pm$6.8) % for the Keio data. The SVM binary classifier achieved a sensitivity of 81.5 ($\pm$3.7)% and specificity of 71.9 ($\pm$5.2)% for the EyePACS validation data, while its corresponding performances for the Keio data were 90.8 ($\pm$5.2)% and 80.0 ($\pm$8.8)%, respectively. Fig. 3 shows the receiver operating characteristic curves for the EyePACS (blue line) and Keio (orange line) data. The AUCs for the two datasets were determined to be 0.868 ($\pm$0.026) and 0.927 ($\pm$0.036), respectively.

Table 1
Baseline statistics of the images.

Characteristic	EyePACS Training dataset (%) (American)	EyePACS Validation dataset (%) (American)	Keio dataset (%) (Japanese)
Total no. of images	2310	720	200
No. of images showing diabetic retinopathy	462 (20.0)	144 (20.0)	40 (20.0)
No. of images showing mild diabetic retinopathy	462 (20.0)	144 (20.0)	40 (20.0)
Not of images showing moderate diabetic retinopathy	462 (20.0)	144 (20.0)	40 (20.0)
No. of images showing severe diabetic retinopathy	462 (20.0)	144 (20.0)	40 (20.0)
No. of images showing Proliferative diabetic retinopathy	462 (20.0)	144 (20.0)	40 (20.0)

The performances of the NN classifiers are shown in Figs. 4 and 5 and Table 3. The total accuracy of the NN classifier was 45.3 ($\pm$3.6)% for the EyePACS validation data, and 38.5 ($\pm$6.7)% for the Keio data. The NN binary classifier achieved a sensitivity of 88.0 ($\pm$3.1)% and specificity of 69.1 ($\pm$5.3)% for the EyePACS data, while its corresponding performances for the Keio data were 87.5 ($\pm$5.9)% and 78.8 ($\pm$9.0)%, respectively. Fig. 5 shows the receiver operating characteristic curves for the EyePACS (blue line) and Keio (orange line) data. The AUCs for the two datasets were determined to be 0.880 ($\pm$0.024) and 0.898 ($\pm$0.043), respectively.

Fig. 6 shows representative images that were misclassified by the SVM classifier. The differences between the images in Fig. 6A and B are due to the data size difference resulting from the different retinal cameras used to capture the images.

The Grad-CAMs of binary classification model are shown in Fig. 7. The figure indicates that the AI was capable of capturing the typical signs of referable DR such as hemorrhage and exudates in both American and Japanese fundus photographs.

4. Discussion

The results obtained by the AI model developed using American fundus images showed that the model could be applied to fundus images for another racial population such as Japanese subjects. However, the accuracy of the five-class classification for the American EyePACS images were better than that for the Japanese Keio images. Nevertheless, the accuracies for both cases were not sufficiently high for a meaningful comparison, owing to the significantly lower classification accuracy for adjacent classes. Conversely, the performance of the binary model for the Keio data was higher than that for the EyePACS data, especially in terms of the ROC-AUC. Retina images differ between ethnic groups and it is important to understand the variations during examination for abnormality [14,15]. Rochtchina et al. noted that the fundus differences between Asian and Caucasian subjects were mainly due to their differing pigment levels [15]. Japanese also exhibit genetic characteristics that are intermediate between Caucasians and Africans [16,17]. It is therefore

A

EyePACS data set	Labels	Predicted Labels 0	1	2	3	4
True Labels	0	67	48	16	7	6
	1	45	58	32	4	5
	2	29	39	43	22	11
	3	10	11	25	75	23
	4	2	5	10	19	108

B

Keio data set (Japanese)	Labels	Predicted Labels 0	1	2	3	4
True Labels	0	21	15	3	1	0
	1	15	12	9	2	2
	2	11	2	7	14	6
	3	1	2	10	25	2
	4	2	2	0	21	15

C

EyePACS data set	Labels	Predicted Labels Non-referable	Referable
True Labels	Non-referable	207	81
	Referable	80	352

D

Keio data set (Japanese)	Labels	Predicted Labels Non-referable	Referable
True Labels	Non-referable	64	16
	Referable	11	109

Fig. 2. Confusion matrices of the five-class classification model for the (A) EyePACS and (B) Keio data, and of the binary classification model for the (C) EyePACS and (D) Keio data. The definitions of the numeric labels are as follows: 0: no DR, 1: mild DR, 2: moderate DR, 3: severe DR, 4: proliferative DR (PDR). The dotted line indicates the boundary for referable DR.

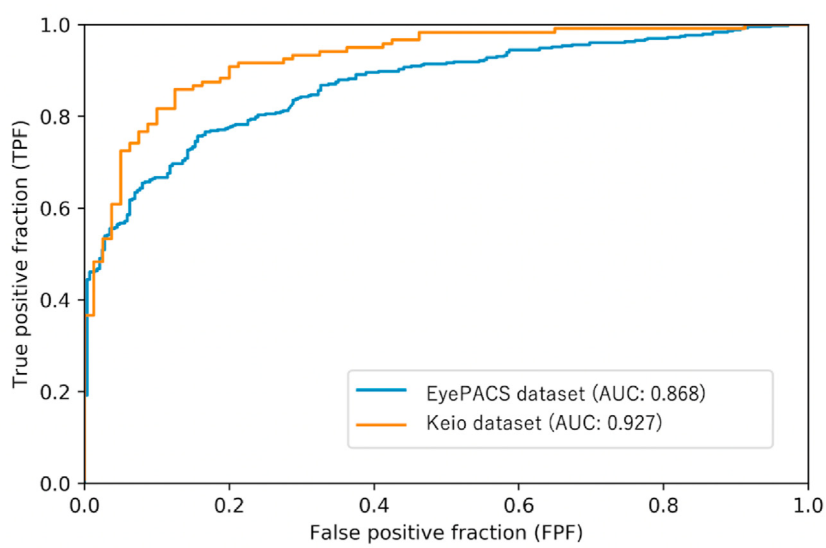

Fig. 3. Receiver operating characteristic curves of the SVM binary classifier for the EyePACS (blue line) and Keio (orange line) data. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2
Classification metrics (SVM).

Data	Grading (five-class model)	Referable DR (binary model)		
	Accuracy	Sensitivity	Specificity	ROC-AUC
EyePACS (American)	48.8 (±3.7) %	81.5 (±3.7) %	71.9 (±5.2) %	0.868 (±0.026)
Keio (Japanese)	40.0 (±6.8) %	90.8 (±5.2) %	80.0 (±8.8) %	0.927 (±0.036)

possible that a system trained on a multiracial dataset would perform better for a Japanese data, or might be sufficiently robust to enable comprehensive adaptation to Japanese. The differing performances of the present AI model for American and Japanese subjects may be attributed to the more distinct fundus differences between the two races, the retinal camera and imaging differences, and the significantly different data sizes. The model can be improved to overcome the inter-racial differences such as the reddish color tone of the Japanese fundus, which can make it more difficult to detect vascular lesions including bleeding. There was no tendency of the SVM nor NN models to under-estimate or overestimate the Japanese fundus (Figs. 2 and 4). Among the

A

EyePACS data set		Predicted Labels				
	Labels	0	1	2	3	4
True Labels	0	52	59	16	10	7
	1	50	54	29	7	4
	2	36	25	40	32	11
	3	10	7	19	85	23
	4	8	1	6	34	95

B

Keio data set (Japanese)		Predicted Labels				
	Labels	0	1	2	3	4
True Labels	0	7	30	1	0	2
	1	12	22	2	3	1
	2	8	9	5	14	4
	3	3	3	4	27	3
	4	0	1	0	23	16

C

EyePACS data set		Predicted Labels	
	Labels	Non-referable	Referable
True Labels	Non-referable	199	89
	Referable	52	380

D

Keio data set (Japanese)		Predicted Labels	
	Labels	Non-referable	Referable
True Labels	Non-referable	63	17
	Referable	15	105

Fig. 4. Confusion matrices of the five-class classification model for the (A) EyePACS and (B) Keio data, and of the binary classification model for the (C) EyePACS and (D) Keio data. The definitions of the numeric labels are as follows: 0: no DR, 1: mild DR, 2: moderate DR, 3: severe DR, 4: proliferative DR (PDR). The dotted line indicates the boundary for referable DR.

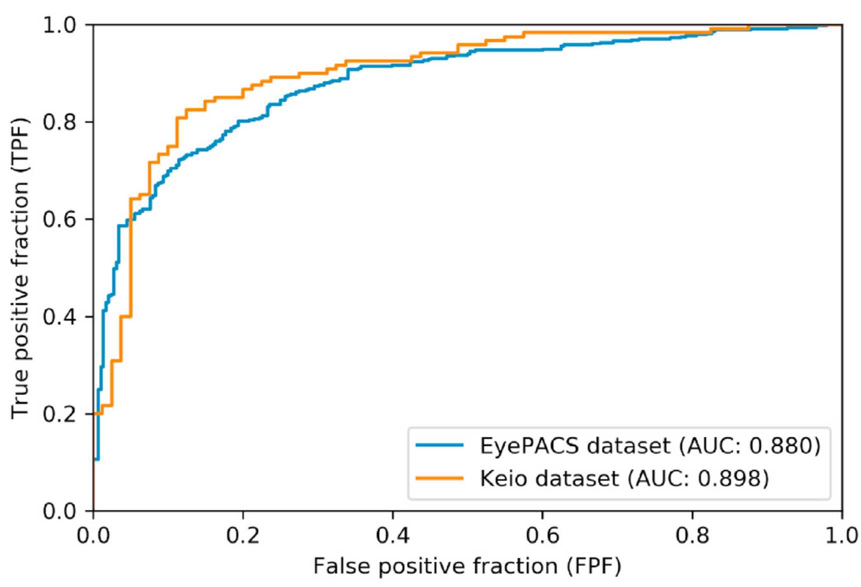

Fig. 5. Receiver Operating Characteristic Curve of the NN binary classifier on EyePACS dataset (blue line) and Keio dataset (orange line). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Classification metrics (NN).

Data	Grading (five-class model)	Referable DR (binary model)		
	Accuracy	Sensitivity	Specificity	ROC-AUC
EyePACS (American)	45.3 (±3.6) %	88.0 (±3.1) %	69.1 (±5.3) %	0.880 (±0.024)
Keio (Japanese)	38.5 (±6.7) %	87.5 (±5.9) %	78.8 (±9.0) %	0.898 (±0.043)

cases in which the AI made a misclassification, there were many images that were difficult to judge even by an ophthalmologist (Fig. 6). When the AI model overestimated the severity, it apparently detected other anomalies such as lens dust, drusen, vitreous opacity, or a tortuous vein. Similar misjudgments often occur among ophthalmologists. Conversely, when the model made an underestimation, it was often difficult to determine the reason. There was no characteristic misdiagnosis tendency between the Japanese and Americans images.

Similar automated DR evaluations have been conducted by other researchers. Gulshan et al. reported an algorithm with a sensitivity of

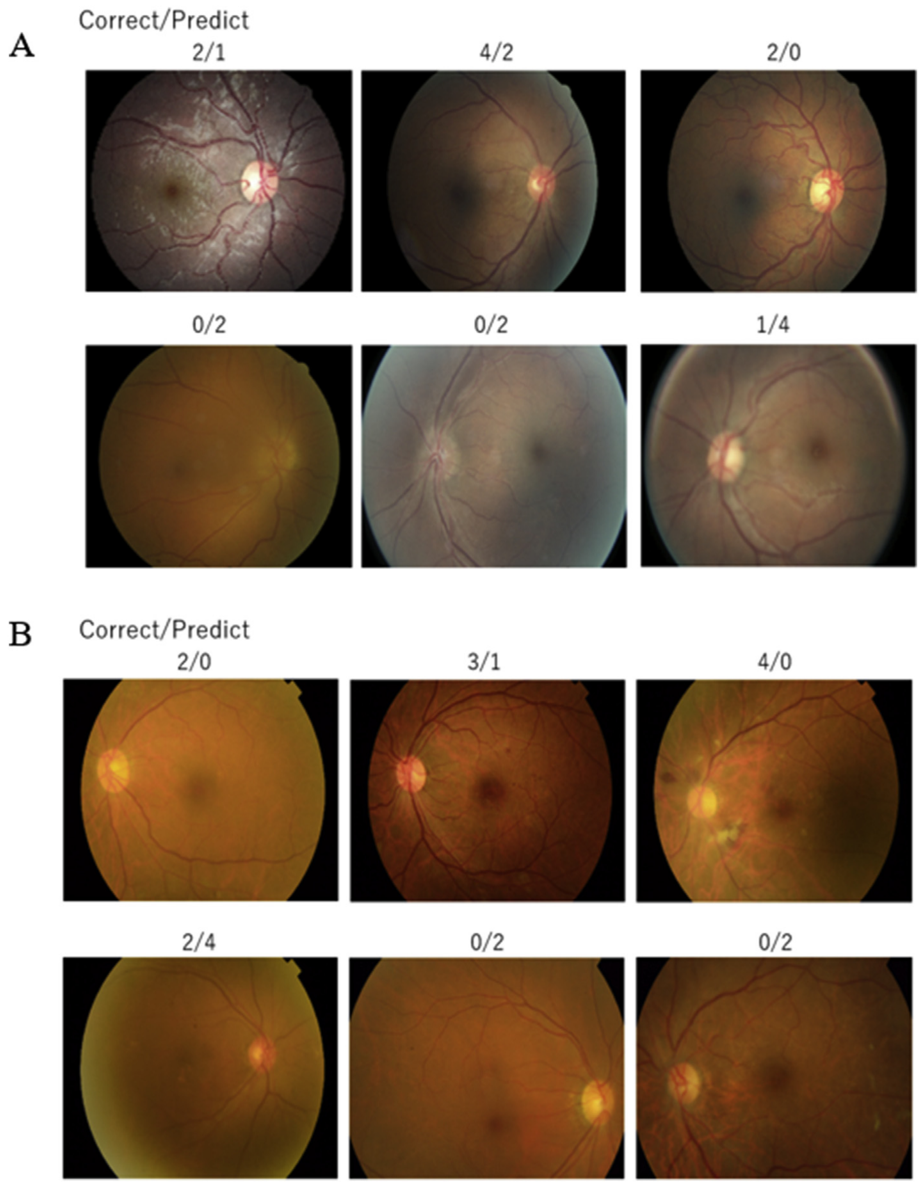

Fig. 6. Examples of (A) EyePACS and (B) Keio images that were misclassified by AI. The numbers indicate the correct/incorrect classification.

90.3% and specificity of 98.1% for detecting referable DR in EyePACS-1 data, based on retrospective training on 128,175 retinal images [3]. Abràmoff et al. reported a similar algorithm with a sensitivity of 96.8% and specificity of 59.4% based on training on publicly available Messidor-2 data [9], while training on 1748 images produced a sensitivity of 97% and specificity of 87% [18]. Philip et al. reported a sensitivity of 86.2% and specificity of 76.8% of an AI method for determining the presence or otherwise of eye disease based on their own dataset consisting of 14,406 images [19]. Lin et al. reported general ophthalmologist-graded referable DR with a sensitivity of 78% and specificity of 86% [20]. It can be seen from the foregoing that most available AI methods can be used for DR screening based on Food and Drug Administration-defined endpoints with a performance level and speed exceeding those of clinicians. In these previous studies, the selection and screening of the learning data were conducted by multiple ophthalmologists. Thus, by increasing the learning data through the addition of Japanese images, improvement of the present model can be expected.

In this study, the SVM architecture exhibited a superior performance compared with the NN architecture. Comparison of SVMs and NNs in previous studies has shown that their equivalent optimal architectures can only be determined on a case-by-case basis [[21–30]]. The sample size can be considered as a factor because the differences are greater for five-class models. While the high accuracies of the previous DR classifications mentioned above were mainly achieved by NN algorithms, the number of samples and other metadata were the main determinants rather than the architecture.

This present work is the first to show that an AI based on US data can be used to grade Japanese fundus photographs. Such an AI system can be used to assist the clinical diagnosis of diabetes and in telemedicine systems. Moreover, the present results indicate the potential of the system for application to other racial populations including in countries with limited medical resources where it is difficult to collect data for AI training. This will contribute to the reduction of the risk of blindness. Although some AI systems output images with insufficient quality, the output of the present system for an image is never "ungradable." Among the 3980 EyePACS images, 472 (11.9%) of insufficient quality were manually excluded, and the Keio images were also collected through screening by trained technicians to ensure that they were gradable.

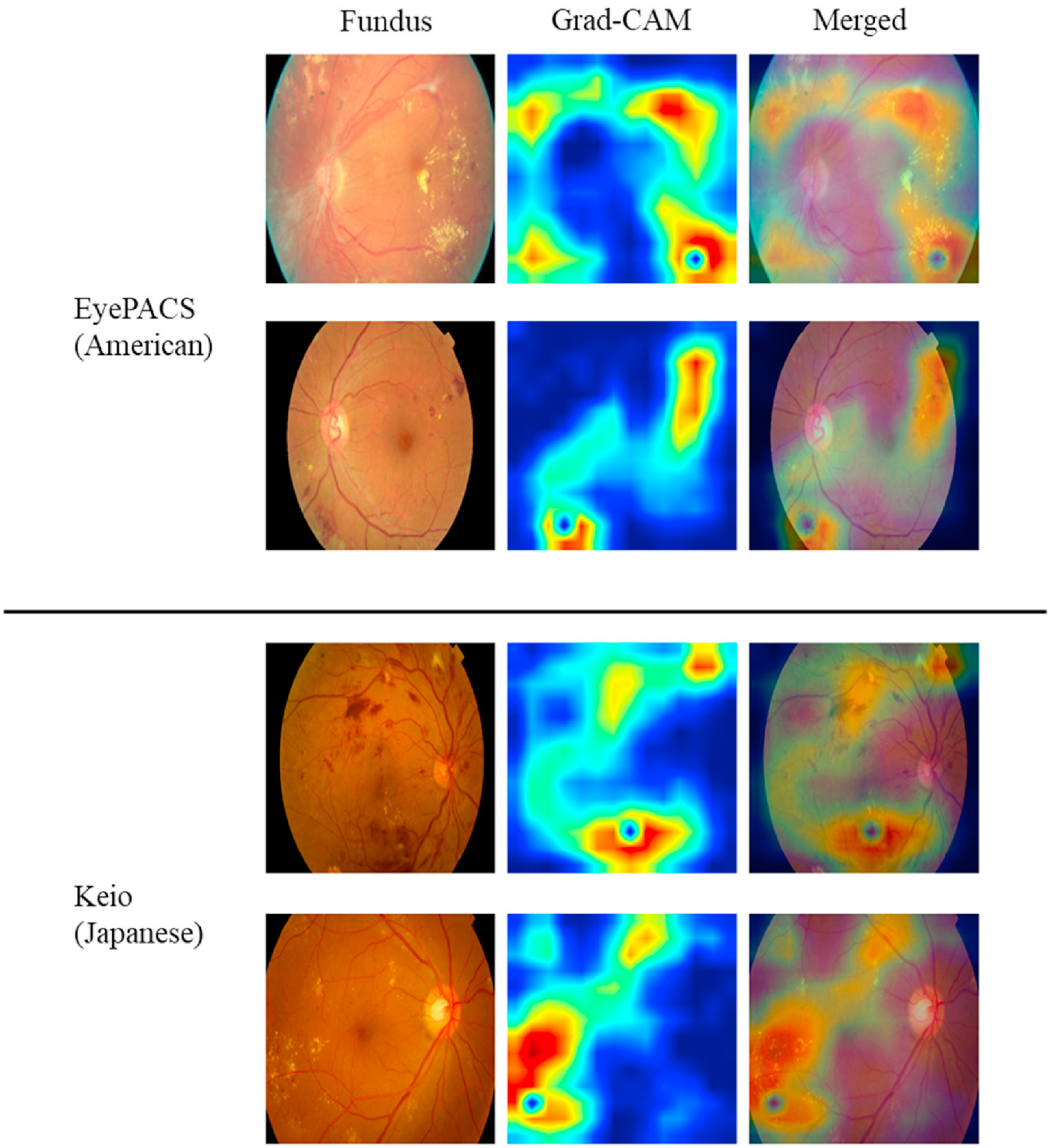

Fig. 7. Examples of the preprocessed fundus photographs and Grad-CAM of NN binary classification model. All images were labeled as referable and also predicted as referable by the model. The aspect ratio is transformed to 1:1 in the preprocessing process.

The proposed system has some limitations. The accuracy is a little too low for clinical application, and the distribution of the patient severity and other background may vary with the actual clinical situation. In the present experiment, the amount of Japanese validation data was small, and it is necessary to conduct further study using larger amounts of both learning and validation data. Further, while the Japanese are almost a single ethnic group, EyePACS consists of multi-ethnic data of unknown racial constitution, and this limits the deductions that can be made from a direct comparison with the Keio. Although EyePACS may include some Japanese data, Japanese Americans represented only 0.4% of the total United States population in the 2018 census [31], and are unlikely to have any relevant effect on the data.

Finally, as reported by Kouroupis et al. color fundus photography does not provide information about the retinal layer structure and morphology. Additional data for the retinal layer can be sourced by optical coherent tomography (OCT), the incorporation of which is expected to improve the classification accuracy of the proposed AI system. This will be implemented in a future work.

5. Conclusions

In this study, the DR classification model using AI created based on American fundus data was able to obtain the better performance even when tested with another racial Japanese population data in detecting referable DR (binary model). Although more experiments were needed to be rigorously proven, it was suggested that the influence of different races has little effect on the fundus grading by the AI model.

Author contributions

Conceptualization: Y.K. and T.K.; methodology: Y.K. and K.M; software: K.M.; validation: Y.K. and N.O.; formal analysis: M.K.; investigation: Y.K. and N.O.; data curation: Y.K., N.O., and Y.O.; writing—original draft preparation: Y.K.; writing—review and editing: T.K.; visualization: K.M.; supervision: K.T. and T.K.; project administration: Y.K., N.O., and T.K.; funding acquisition: Y.K.

Funding

Y.K. got funded by 4th Grants 4 Apps Tokyo held by Bayer AG, Germany.

Declaration of interest

Y.K. obtained funding from 4th Grants 4 Apps Tokyo held by Bayer AG. The funder was not involved in the design of the study, the collection, analyses, and interpretation of data, the writing of the manuscript, nor the decision to publish the results. The other authors declare no conflict of interest.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank S. Ikeda, A. Ishida, Y. Miwa, A. Shinojima, C, Shoda, Y. Hagiwara, and Y. Nitsuma of the Laboratory of Photobiology, Keio University School of Medicine for their technical and administrative support. Funding from Bayer AG, Germany is gratefully acknowledged.

References

[1] Morizane Y, Morimoto N, Fujiwara A, Kawasaki R, Yamashita H, Ogura Y, Shiraga F. Incidence and causes of visual impairment in Japan: the first nation-wide complete enumeration survey of newly certified visually impaired individuals. Jpn J Ophthalmol 2019;63(1):26–33. https://doi.org/10.1007/s10384-018-0623-4.

[2] Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. Ophthalmology 2017;124(7):962–9. https://doi.org/10.1016/j.ophtha.2017.02.008.

[3] 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. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016; 304(6):649–56. https://doi.org/10.1001/jama.2016.17216.

[4] Gulshan V, Rajan RP, Widner K, Wu D, Wubbels P, Rhodes T, Whitehouse K, Coram M, Corrado G, Ramasamy K, Raman R, Peng L, Webster DR. Performance of a deep-learning algorithm vs manual grading for detecting diabetic retinopathy in India. JAMA Ophthalmol. 2019;137(9):987–93. https://doi.org/10.1001/jamaophthalmol.2019.2004.

[5] Ruamviboonsuk P, Krause J, Chotcomwongse P, Sayres R, Raman R, Widner K, Campana BJL, Phene S, Hemarat K, Tadarati M, Silpa-Archa S, Limwattanayingyong J, Rao C, Kuruvilla O, Jung J, Tan J, Orprayoon S, Kangwanwongpaisan C, Sukumalpaiboon R, Webster DR. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. NPJ Digital Med 2019;2(1). https://doi.org/10.1038/s41746-019-0099-8.

[6] Sayres R, Taly A, Rahimy E, Blumer K, Coz D, Hammel N, Krause J, Narayanaswamy A, Rastegar Z, Wu D, Xu S, Barb S, Joseph A, Shumski M, Smith J, Sood AB, Corrado GS, Peng L, Webster DR. Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. Ophthalmology 2019;126(4):552–64. https://doi.org/10.1016/j.ophtha.2018.11.016.

[7] Ting DSW, Cheung CYL, Lim G, Tan GSW, Quang ND, Gan A, Hamzah H, Garcia-Franco R, Yeo IYS, Lee SY, Wong EYM, Sabanayagam C, Baskaran M, Ibrahim F, Tan NC, Finkelstein EA, Lamoureux EL, Wong IY, Bressler NM, Wong TY. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA - J Am Med Assoc 2017;318(22):2211–23. https://doi.org/10.1001/jama.2017.18152.

[8] n.d. American Academy of Ophthalmology. International Council of Ophthalmology : International Clinical Diabetic Retinopathy Disease Severity Scale. Detailed Table. Retrieved August 12, 2019, from, http://www.icoph.org/dynamic/attachments/resources/diabetic-retinopathy-detail.pdf.

[9] Abràmoff Michael D, Folk JC, Han DP, Walker JD, Williams DF, Russell SR, Massin P, Cochener B, Gain P, Tang L, Lamard M, Moga DC, Quellec G, Niemeijer M. Automated analysis of retinal images for detection of referable diabetic retinopathy. JAMA Ophthalmol 2013;131(3):351. https://doi.org/10.1001/jamaophthalmol.2013.1743.

[10] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L. ImageNet large scale visual recognition challenge. Int J Comput Vis 2015;115(3):211–52. https://doi.org/10.1007/s11263-015-0816-y.

[11] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 2818–26. https://doi.org/10.1109/CVPR.2016.308.

[12] Platt J, others. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Adv Large Margin Classifiers 1999;10(3):61–74.

[13] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. ArXiv 2016. https://doi.org/10.1007/s11263-019-01228-7.

[14] Bourne RRA. Ethnicity and ocular imaging. Eye 2011;25(3):297–300. https://doi.org/10.1038/eye.2010.187.

[15] Rochtchina E, Wang JJ, Taylor B, Wong TY, Mitchell P. Ethnic variability in retinal vessel caliber: a potential source of measurement error from ocular pigmentation?-The Sydney childhood eye study. Invest Ophthalmol Vis Sci 2008;49(4):1362–6. https://doi.org/10.1167/iovs.07-0150.

[16] Edwards M. The genetic architecture of Iris colour and surface feature variation in populations of diverse ancestry. 2016.

[17] Spichenok O, Budimlija ZM, Mitchell AA, Jenny A, Kovacevic L, Marjanovic D, Caragine T, Prinz M, Wurmbach E. Prediction of eye and skin color in diverse populations using seven SNPs. Forensic Sci Int: Genetics 2011;5(5):472–8. https://doi.org/10.1016/j.fsigen.2010.10.005.

[18] Abràmoff Michael David, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Opthalmol Vis Sci 2016;57(13):5200. https://doi.org/10.1167/iovs.16-19964.

[19] Philip S, Fleming AD, Goatman KA, Fonseca S, Mcnamee P, Scotland GS, Prescott GJ, Sharp PF, Olson JA. The efficacy of automated “disease/no disease” grading for diabetic retinopathy in a systematic screening programme. Br J Ophthalmol 2007;91(11):1512–7. https://doi.org/10.1136/bjo.2007.119453.

[20] Lin DY, Blumenkranz MS, Brothers RJ, Grosvenor DM. The sensitivity and specificity of single-field nonmydriatic monochromatic digital fundus photography with remote image interpretation for diabetic retinopathy screening: a comparison with ophthalmoscopy and standardized mydriatic color photography11Inter. Am J Ophthalmol 2002;134(2):204–13. https://doi.org/10.1016/S0002-9394(02)01522-2.

[21] Agarap AF. An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification. http://arxiv.org/abs/1712.03541; 2017.

[22] Choudhary M, Tiwari V, Venkanna U. An approach for iris contact lens detection and classification using ensemble of customized DenseNet and SVM. Future Generat Comput Syst 2019;101:1259–70. https://doi.org/10.1016/j.future.2019.07.003.

[23] Dang LM, Hassan SI, Suhyeon I, Sangaiah AK, Mehmood I, Rho S, et al. UAV based wilt detection system via convolutional neural networks. Sustain Comput: Informat Syst, October 2018;2017:100250. https://doi.org/10.1016/j.suscom.2018.05.010.

[24] Dong Y, Zhang Q, Qiao Z, Yang JJ. Classification of cataract fundus image based on deep learning. In: IST 2017-IEEE International Conference on Imaging Systems and Techniques, Proceedings; 2017. p. 1–5. https://doi.org/10.1109/IST.2017.8261463. 2018-Janua.

[25] Lang R, Jia K, Feng J. Brain tumor identification based on CNN-SVM model. In: ACM International Conference Proceeding Series; 2018. p. 31–5. https://doi.org/10.1145/3278198.3278209.

[26] Patalas-Maliszewska J, Halikowski D. A model for generating workplace procedures using a CNN-SVM architecture. Symmetry 2019;11(9):1–15. https://doi.org/10.3390/SYM11091151.

[27] Peng Y, Liao M, Deng H, Ao L, Song Y, Huang W, Hua J. CNN-SVM: a classification method for fruit fL image with the complex background. IET Cyber-Phys Syst: Theor Appl 2020;5(2):1–5. https://doi.org/10.1049/iet-cps.2019.0069.

[28] Qi X, Wang T, Liu J. Comparison of support vector machine and softmax classifiers in computer vision. In: Proceedings - 2017 2nd International Conference on Mechanical, Control and Computer Engineering, ICMCCE 2017; 2018. p. 151–5. https://doi.org/10.1109/ICMCCE.2017.49. 2018-Janua.

[29] Tang Y. Deep learning using linear support vector machines. Icml 2013. http://arxiv.org/abs/1306.0239.

[30] Xue DX, Zhang R, Feng H, Wang YL. CNN-SVM for microvascular morphological type recognition with data augmentation. J Med Biol Eng 2016;36(6):755–64. https://doi.org/10.1007/s40846-016-0182-4.

[31] 2018 Census Test. United States census bureau. n.d. https://data.census.gov/cedsci/.