Accuracy of deep learning, a machine learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting idiopathic macular holes

Toshihiko Nagasawa [Corresp., 1], **Hitoshi Tabuchi** [1], **Hiroki Masumoto** [1], **Hiroki Enno** [2], **Masanori Niki** [3], **Hideharu Ohsugi** [1], **Yoshinori Mitamura** [3]

[1] Department of Ophthalmology, Tsukazaki Hospital, Himeji City, Hyogo Prefecture, Japan

[2] Rist Inc., Tokyo, Japan

[3] Department of Ophthalmology, Institute of Biomedical Sciences, Tokushima University, Tokushima City, Tokushima Prefecture, Japan

Corresponding Author: Toshihiko Nagasawa
Email address: t.nagasawa@tsukazaki-eye.net

We aimed to investigate the detection of idiopathic macular holes (MHs) using ultra-wide-field fundus images (Optos) with deep learning, which is a machine learning technology. The study included 910 Optos color images (715 normal images, 195 MH images). Of these 910 images, 637 were learning images (501 normal images, 136 MH images) and 273 were test images (214 normal images and 59 MH images). We conducted training with a deep convolutional neural network (CNN) using the images and constructed a deep-learning model. The CNN exhibited high sensitivity of 100% (95% confidence interval [CI], 93.5–100%) and high specificity of 99.5% (95% CI, 97.1–99.9%). The area under the curve was 0.9993 (95% CI, 0.9993–0.9994). Our findings suggest that MHs could be diagnosed using an approach involving wide angle camera images and deep learning.

Accuracy of deep learning, a machine learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting idiopathic macular holes

Toshihiko Nagasawa,[1] Hitoshi Tabuchi,[1] Hiroki Masumoto,[1] Hiroki Enno,[2] Masanori Niki,[3] Hideharu Ohsugi,[1] Yoshinori Mitamura[3]

1 Department of Ophthalmology, Tsukazaki Hospital, Himeji, Japan

2 Rist Inc., Tokyo, Japan

3 Department of Ophthalmology, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, Japan

Corresponding Author:

Toshihiko Nagasawa[1]

68-1 Aboshi Waku, Himeji City, Hyogo Prefecture 671-1227, Japan;

15 Tel: +81 79-272-8555; Fax: +81 79-272-8550

16 E-mail: t.nagasawa@tsukazaki-eye.net

17

18 Abstract

19 We aimed to investigate the detection of idiopathic macular holes (MHs) using ultra-wide-field
20 fundus images (Optos) with deep learning, which is a machine learning technology. The study
21 included 910 Optos color images (715 normal images, 195 MH images). Of these 910 images,
22 637 were learning images (501 normal images, 136 MH images) and 273 were test images (214
23 normal images and 59 MH images). We conducted training with a deep convolutional neural
24 network (CNN) using the images and constructed a deep-learning model. The CNN exhibited
25 high sensitivity of 100% (95% confidence interval [CI], 93.5–100%) and high specificity of
26 99.5% (95% CI, 97.1–99.9%). The area under the curve was 0.9993 (95% CI, 0.9993–0.9994).
27 Our findings suggest that MHs could be diagnosed using an approach involving wide angle
28 camera images and deep learning.

29

30 Introduction

31 In 1988, Gass described idiopathic macular holes (MHs) as a retinal break commonly involving
32 the fovea (Gass, 1998), and in 1991 Kelly and Wendel reported that MHs can be successfully
33 repaired through vitreous surgery (Kelly & Wendel, 1991). The age and gender adjusted annual
34 incidences of primary MH have been reported at 7.9 eyes and 7.4 respectively per 100 000
35 inhabitants, and the male to female ratio was 1:2.2 (Forsaa et al., 2017). The accepted
36 pathogenesis has macular hole formation proceeding in stages from an impending hole to a full
37 thickness MH, with visual acuity deteriorating to less than 6/60 in 85% of cases (Luckie &
38 Heriot, 1995). The development of optical coherence tomography (OCT) and improvement of
39 image resolution have made the diagnosis of macular diseases substantially easy (Kishi &
40 Takahashi, 2000).

41 In addition, the advent of wide angle fundus cameras has made the observation of the entire
42 retina possible through a simple and noninvasive approach (Nagiel et al., 2016). An example of

such a camera is the ultra-wide-field scanning laser ophthalmoscope (Optos 200 Tx; Optos PLC, Dunfermline, United Kingdom), which is known as Optos. It is capable of photographing the fundus without mydriasis, and it is used for making judgments regarding the diagnosis, follow-up, and treatment effects of various fundus diseases (Prasad et al., 2010; Wessel et al., 2012; Ogura et al., 2014). Optos can minimize the risk of a rise in pupillary block caused by mydriasis and intraocular pressure increase. This makes Optos suitable for medical use in remote areas where the services of ophthalmologists are limited, as the device can be safely used by orthoptists and other medical professionals.

Recently, image processing technology applying deep learning, a sub-field of machine learning algorithm studies, has attracted attention because of its very high classification performance. The use of this technology for medical images is being actively studied (LeCun, Bengio & Hinton, 2015; Liu et al., 2015; Litjens et al., 2016). In the ophthalmic field, there are reports on the use of the ocular fundus camera and deep learning and on the improvement in the accuracy of automatic diagnosis of diabetic retinopathy and retinal detachment with these

57 approaches (Gulshan et al., 2016; Ohsugi et al., 2017; Ryan et al., 2018). However, the
58 diagnostic accuracy of the wide angle ocular fundus camera for macular diseases is yet to be
59 reported. Deep neural networks have been used to diagnose skin cancer with as much accuracy
60 as that attained by dermatologists (Esteva et al., 2017). We decided to assess the diagnostic
61 capability of deep neural networks for macular holes as compared with ophthalmologists'
62 diagnoses.

63 The present study assessed the presence of MHs, which are considered as a macular disease,
64 using ultra-wide-field fundus images with deep learning in order to determine the accuracy of
65 deep learning, and to compare the ophthalmologist and the deep neural network for MHs.

66

67 Materials and Methods

68 Data set

69 The study dataset included 910 Optos color images obtained at the Tsukazaki Hospital (Himeji,
70 Japan) and Tokushima University Hospital (715 normal images and 195 MH images). Of the 910

71 images, 637 were used for training purposes (80%; 501 normal images and 136 MH images;
72 learning images) and 273 were used for testing purposes (20%; 214 normal images and 59 MH
73 images; test images).

74 The 637 learning images underwent image processing and were amplified to 5000 images
75 (3887 normal images and 1113 MH images). The image amplification process comprised
76 contrast adjustment, γ correction, histogram equalization, noise addition, and inversion. We
77 performed training on these learning images with a deep convolutional neural network (CNN)
78 and constructed a deep learning model.

79 Cases of MHs were confirmed by a retinal specialist who conducted fundus examinations
80 using an ophthalmoscope and OCT. For OCT, a swept-source OCT system (SS-OCT; DRI OCT-
81 1 Atlantis, TOPCON Corporation, Tokyo, Japan) was used. All Optos images obtained from the
82 MH patient database were considered for inclusion. Images from patients complications, such as
83 vitreous hemorrhage, asteroid hyalosis, intense cataract, and retinal photocoagulation scars, and
84 other conditions, such as fundus diseases, were excluded. Additionally, images with poor clarity

were excluded. Moreover, images from patients with stage 1 MHs (according to the classification by Gass) and those with retinal detachment were excluded.

The procedures used conformed to the tenets of the Declaration of Helsinki, and an informed consent was obtained from either the subjects or their legal guardians after explanation of the nature and possible consequences of the study. An approval was obtained from the Institutional Review Board of Tsukazaki Hospital (No 171001) and Tokushima University Hospital (No 3079) to perform this study.

Deep learning model

We implemented a deep learning model using a CNN (Figure 1). We arranged three convolutional layers. The rectified linear unit (ReLU) activation function and batch normalization were placed after each convolutional layer. A max pooling layer (MP 1, 2) was placed after convolutional layers 1 and 3. In addition, a dropout layer (drop rate 0.25) was placed after each max pooling layer (MP 1, 2). Finally, the two fully connected layers (FC 1, 2) were

99 arranged and classified into two classes using the Softmax function.

100

101 Training the deep convolutional neural network

102 All obtained image data were converted to 256 × 192 pixels. Learning was carried out with

103 mini-batch processing of 10 images and an epoch number of 100. The initial value of the

104 network weight was randomly provided as the zero average of Gaussian distribution, with a

105 standard deviation of 0.05. Dropout processing was performed to mask the first total tie layer

106 (FC1), with 50% probability. The network weights were optimized using stochastic gradient

107 descent (SGD) with momentum [learning coefficient, 0.01; inertia term, 0.9]. Of 100 deep

108 learning models obtained in 100 learning cycles, the model with the highest accuracy rate for the

109 test data was selected as the deep learning model.

110

111 Outcome

112 The area under the curve (AUC) and sensitivity/specificity were determined for the ability of the

selected CNN model to discriminate between normal eyes and MH.

Statistical analysis

The receiver operating characteristic curve (ROC curve) and the 95% confidence interval (CI) of the AUC were obtained. The ROC curve was created by considering that the value judged to involve MHs exceeded the threshold (cutoff value) as positive. The model was fitted to only 90% of the test data, and 10% were thinned out. We created 100 ROC curves by making 100 patterns. One hundred AUCs were calculated from the ROC curves. With regard to the AUCs, 95% CI were obtained by assuming normal distribution and using standard deviation. With regard to sensitivity and specificity, the first of the 100 ROC curves were used, and the sensitivity and specificity at the optimum cutoff value calculated using Youden Index 23 as the representative value of the deep learning model were used. The accuracy, specificity, sensitivity, and response times by CNN and six ophthalmologists were calculated.

127 Creation of an ophthalmologist application

128 Of the 273 test images, 50 normal images and 50 MH images were extracted using the random
129 number generation method (equal representation for normal data and the disease data). We
130 calculated the accuracy, specificity, sensitivity, and response times by CNN based on the
131 averaged results of six ophthalmologists.

132

133 Determination and measurement methods for calculating the required time

134 Six ophthalmologists determined the presence or absence of MHs in 50 images presented on a
135 computer monitor. The answer inputs of either 0 or 1 on the response form were populated in an
136 Excel table.

137 The time taken by the ophthalmologists to enter data in the computer was also included. In deep
138 neural network, a series of tasks was performed for all presented numbers as follows: confirming
139 the number of the problem in the answer column → reading the image → judging → filling in
140 the answer column. The total time was counted as the operation time. This series of work was

141 performed 15 times by a computer, and the working time was considered as the median value.
142 The time required by the ophthalmologists was set as the time taken to complete all answers in
143 the Excel file. The time required for the deep neural network was measured by the internal clock
144 of the computer. The specifications of the computer were as follows: operating system, Windows
145 10 Home; CPU, Intel Core i7 - 3630 QM; memory, 8.00 GB; GPU, NA.
146
147 Heat map
148 Using the gradient weighted class activation mapping (Grad-CAM) (Selvaraju et al., 2017)
149 method, we obtained a heat map of the coordinate axes in the image focused on by the CNN. The
150 layer that used the gradient was specified as convolution layer 2. Additionally, we specified
151 ReLU as the backprop modifier.
152
153 Results
154 Background data

155 Table 1 shows the total number of normal and MH images, patient age, patient sex, and left/right
156 of the imaged eyes. There were no statistically significant differences between the normal and
157 MH images with regard to age, sex ratio, and left eye ratio (Student's *t*-test and Fisher's exact
158 test).
159
160 Evaluation of the performance model
161 The mean value of 100 AUCs prepared by the CNN model was 0.9993 (95%–CI: 0.9993–
162 0.9994).
163 The first curve among the 100 calculated ROC curves is shown in Figure 2.
164 The mean sensitivity obtained from the 100 ROC curves was 100% (95%–CI; 93.5–100%), and
165 the mean specificity was 99.5% (95%–CI; 97.1–99.99%).
166 Ophthalmologists carried out the test, and the mean (standard deviation) required time was
167 838.00 s (±199.16), the mean (standard deviation) accuracy rate was 80.6% (5.9%), sensitivity
168 was 65.9% (15.7%), and specificity was 95.2% (4.3%). The same test was carried out with the

169 CNN model, and the mean (standard deviation) required time was 32.8 s (±7.36) and accuracy
170 rate, sensitivity, and specificity were all 100% (Table 2).
171
172 Heat map
173 An image with the corresponding heat map superimposed was created by the CNN, and the
174 focused coordinate axes in the image were indicated. A representative image is presented in
175 Figure 3. Focal points accumulated on the heat map at the fovea of the fundus macula. It is
176 suggested that the CNN may distinguish s diseased eye from a normal eye by focusing on the
177 MH lesion site.
178 Blue color was used to indicate the strength of CNN attention. The color became stronger on one
179 side of the arcade, with centering at the macular fovea, and accumulation was noted at the focus
180 points.
181
182 Discussion

OCT is considered indispensable for the diagnosis of MHs. However, in the present study, MHs were diagnosed using images from a wide angle camera and deep learning. Optos adopts the method of combining a red (633 nm) laser image and a green (532 nm) laser image to give a false color. Details of color information are inferior to those of a conventional fundus camera. Therefore, the quality of the diagnosis made by an ophthalmologist might reduce. With the deep learning model, the approach is different from the approach of an ophthalmologist, with a focus only on the difference from a normal eye, and there is a possibility that some additional general and flexible features of learning can be considered. The heat map spreads over a relatively wide area around the macula fovea, and this approach appears to have a classification that is superior to the judgment ability of an ophthalmologist.

The present study has several limitations. When light transmission in the eye is absent because of intense cataract or dark vitreous hemorrhage, it is difficult to obtain images with Optos, and such cases were not included in the present study. In addition, this study only compared normal eyes and MH eyes, and it did not assess eyes affected by other fundus diseases. This warrants the

197 preparation of a large scale data set for applying deep learning. Although the diagnostic ability of
198 using a wide angle ocular fundus camera and deep learning for diabetic retinopathy and retinal
199 detachment has been reported, the findings of this study indicate the high diagnostic ability of
200 this approach for MHs, which are considered a macular disease. In the future, studies should
201 assess the possibility of performing automatic diagnoses with a wide angle camera for other
202 macular diseases, such as macular epiretinal membrane and age-related macular degeneration.
203 If Optos is used in a medically depopulated area, wide-area ocular fundus photography can easily
204 be performed under a non-mydriasis condition, without medical complications. Moreover, even
205 if no ophthalmologist is available to assess the image, the deep-learning algorithm can be used
206 for MH diagnosis, as it has a high accuracy rate for MH diagnosis. Many regions of the world
207 have an inadequate number of ophthalmologists (Resnikoff et al., 2012) and thus, the automatic
208 diagnosis of MH using Optos fundus images has great potential. If surgical treatment is
209 performed at an appropriate time in MH patients, a good prognosis can be obtained. The results
210 of this study strongly support the use of an Optos based telemedicine system. Such systems

might aid in the early detection of patients with MHs in areas where ophthalmologists are absent.

Conclusions

Using ultra-wide-field fundus images, deep learning, could successfully diagnose MHs. We believe that this approach will be very useful in the practical clinical diagnosis of MHs. Further research with increasing number of sheets, deepening the layer structure, and using metastasis learning are necessary to confirm our results.

Acknowledgments

The authors would like to thank Enago (www.enago.jp) for the English language review.

References

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:115-118 DOI:

10.1038/nature22985.

Forsaa VA, Lindtjørn B, Kvaløy JT, Frøystein T, Krohn J. (2017) Epidemiology and morphology of full-thickness macular holes. Acta Ophthalmologica DOI:10.1111/aos.13618.

Gass JDM. 1988. Idiopathic senile macular hole: its early stages and pathogenesis. Archives of Ophthalmology 106:629-639 DOI:10.1001/archopht.1988.01060130683026.

Gass JD. 1995. Reappraisal of biomicroscopic classification of stages of development of a macular hole. American Journal of Ophthalmology 119:752-759 DOI:10.1016/S0002-9394(14)72781-3.

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. 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316: 2402-2410 DOI:10.1001/jama.2016.17216.

Kelly NE, Wendel RT. 1991. Vitreous surgery for idiopathic macular holes: results of a pilot study. Archives of Ophthalmology 109:654-649 DOI:10.1001/archopht.1991.01080050068031.

239 Kishi S, Takahashi H. 2000. Three-dimensional observations of developing macular holes.
240 American Journal of Ophthalmology 130:65-75 DOI:10.1016/S0002-9394(00)00383-4.
241 LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. Nature 521:436-444 DOI
242 22910.1038/nature14539.
243 Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-van de
244 Kaa C, Bult P, van Ginneken B, van der Laak J. 2016. Deep learning as a tool for increased
245 accuracy and efficiency of histopathological diagnosis. Scientific Reports 6:26286. DOI:
246 10.1038/srep26286.
247 Liu S, Cai W, Che H, Pujol S, Kikinis R, Feng D, Fulham MJ. 2015. Multimodal neuroimaging
248 feature learning for multiclass diagnosis of Alzheimer's disease. IEEE Transactions on
249 Biomedical Engineering 62:1132-1140 DOI:10.1109/TBME.2014.2372011.
250 Luckie A, Heriot W. 1995. Macular holes. Pathogenesis, natural history and surgical outcomes.
251 Aust N Z J Ophthalmol 23:93-100.
252 Nagiel A, Lalane RA, Sadda, SR, Schwartz SD. 2016. Ultra-widefield fundus imaging: a

253 review of clinical applications and future trends. Retina 36:660-678 DOI
254 10.1097/IAE.0000000000000937.

255 Ogura S, Yasukawa T, Kato A, Usui H, Hirano Y, Yoshida M, Ogura Y. 2014. Wide-field
256 fundus autofluorescence imaging to evaluate retinal function in patients with retinitis pigmentosa.
257 American Journal of Ophthalmology 158:1093-1098 DOI: 10.1016/j.ajo.2014.07.021.

258 Ohsugi H. Tabuchi H, Enno H, Ishitobi N. 2017. Accuracy of deep learning, a machine-learning
259 technology, using ultra–wide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal
260 detachment. Scientific Reports 7:9425. DOI: 10.1038/s41598-017-09891-x.

261 Prasad PS, Oliver SC, Coffee RE, Hubschman JP, Schwartz SD. 2010. Ultra wide-field
262 angiographic characteristics of branch retinal and hemicentral retinal vein occlusion.
263 Ophthalmology 117:780-784 DOI 10.1016/j.ophtha.2009.09.019.

264 Resnikoff S, Felch W, Gauthier TM, Spivey B. 2012. The number of ophthalmologists in
265 practice and training worldwide: a growing gap despite more than 200,000 practitioners. British
266 Journal Ophthalmology 96:783-787 DOI:10.1136/bjophthalmol-2011-301378.

267 Ryan Poplin, Avinash V Varadarajan, Katy Blumer, Yun Liu, Michael V McConnell, Greg S
268 Corrado, Lily Peng, Dale R Webster. 2018. Prediction of cardiovascular risk factors from retinal
269 fundus photographs via deep learning. Nature Biomedical Engineeringvolume 2:158-164
270 DOI:10.1038/s41551-018-0195-0.

271 Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. 2016. Grad-CAM: Visual
272 Explanations from Deep Networks via Gradient-based Localization.
273 https://arxiv.org/abs/1610.02391.

274 Wessel MM, Aaker GD, Parlitsis G, Cho M, D'Amico DJ, Kiss S. 2012. Ultra-wide-field
275 angiography improves the detection and classification of diabetic retinopathy. Retina 32:785-791
276 DOI:10.1097/IAE.0b013e3182278b64.

Figure 1

Overall architecture of the deep learning model.

First, each dataset's image was reduced to 256 × 192 and was input into the model. Next, it was passed through all convolution layers and the entire binding layer, and it was classified into 2 classes.

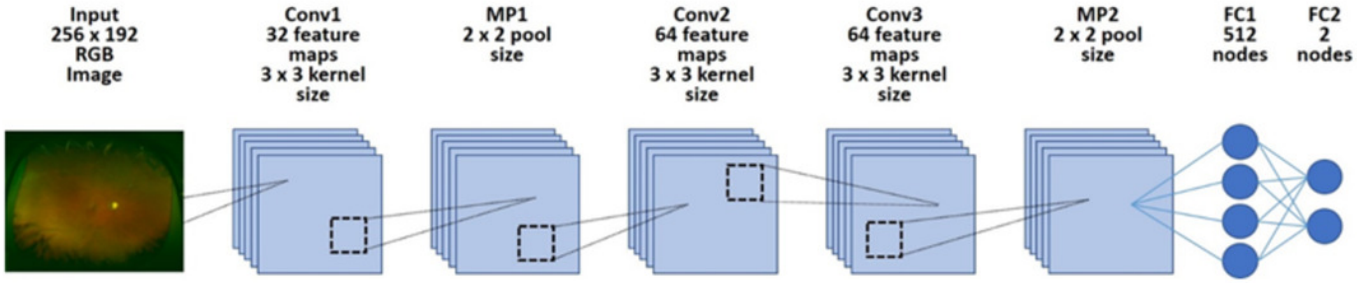

Figure 2

Receiver operating characteristics curve.

This is the first one out of 100 ROC curves. The average AUC of 100 ROC curves was almost 1, and all ROC curves were similar.

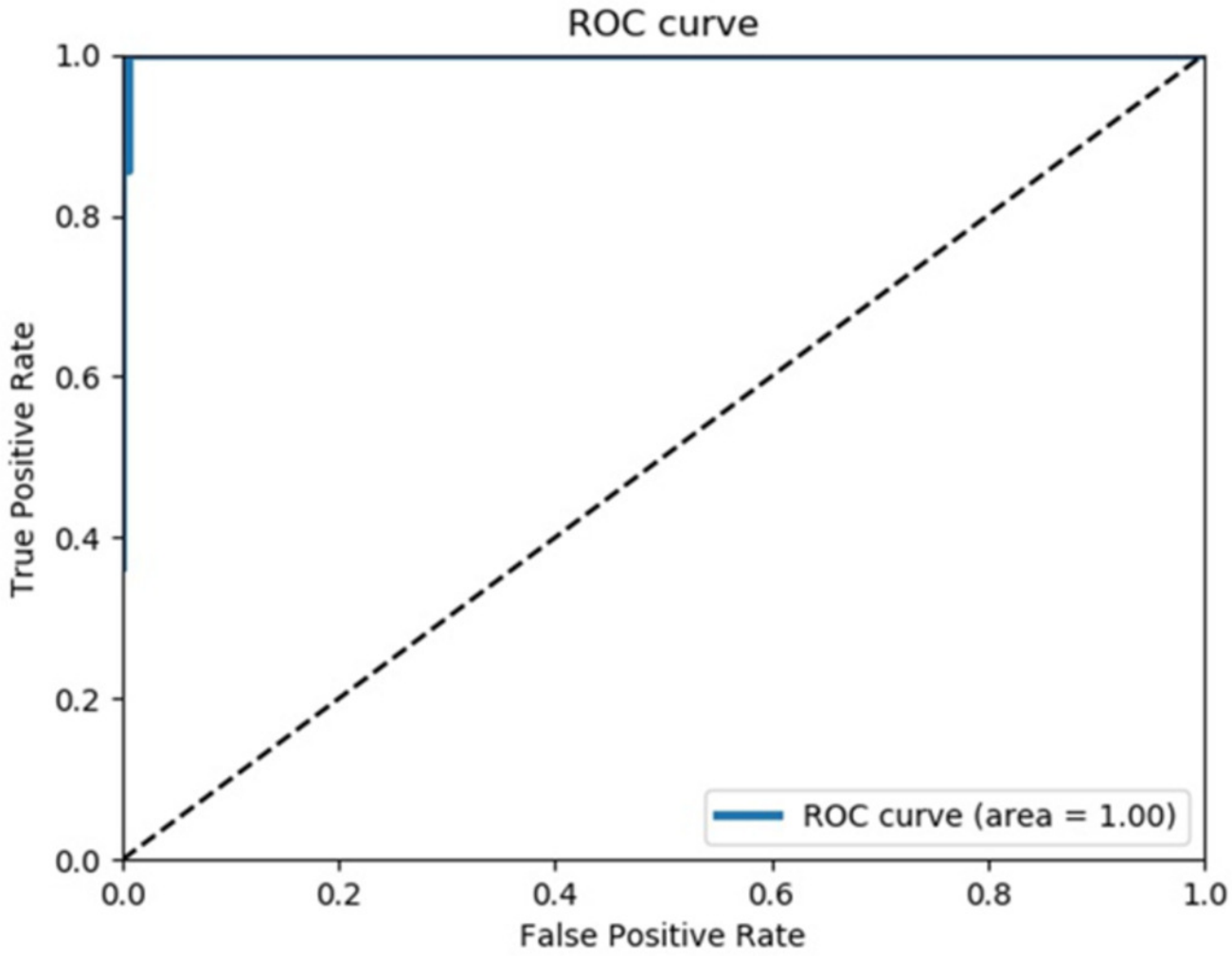

Figure 3

Heatmap superimposed on the photo.

The dark blue color shows the point where the deep neural network is paying attention on the macula and from the same point of view of an ophthalmologist.

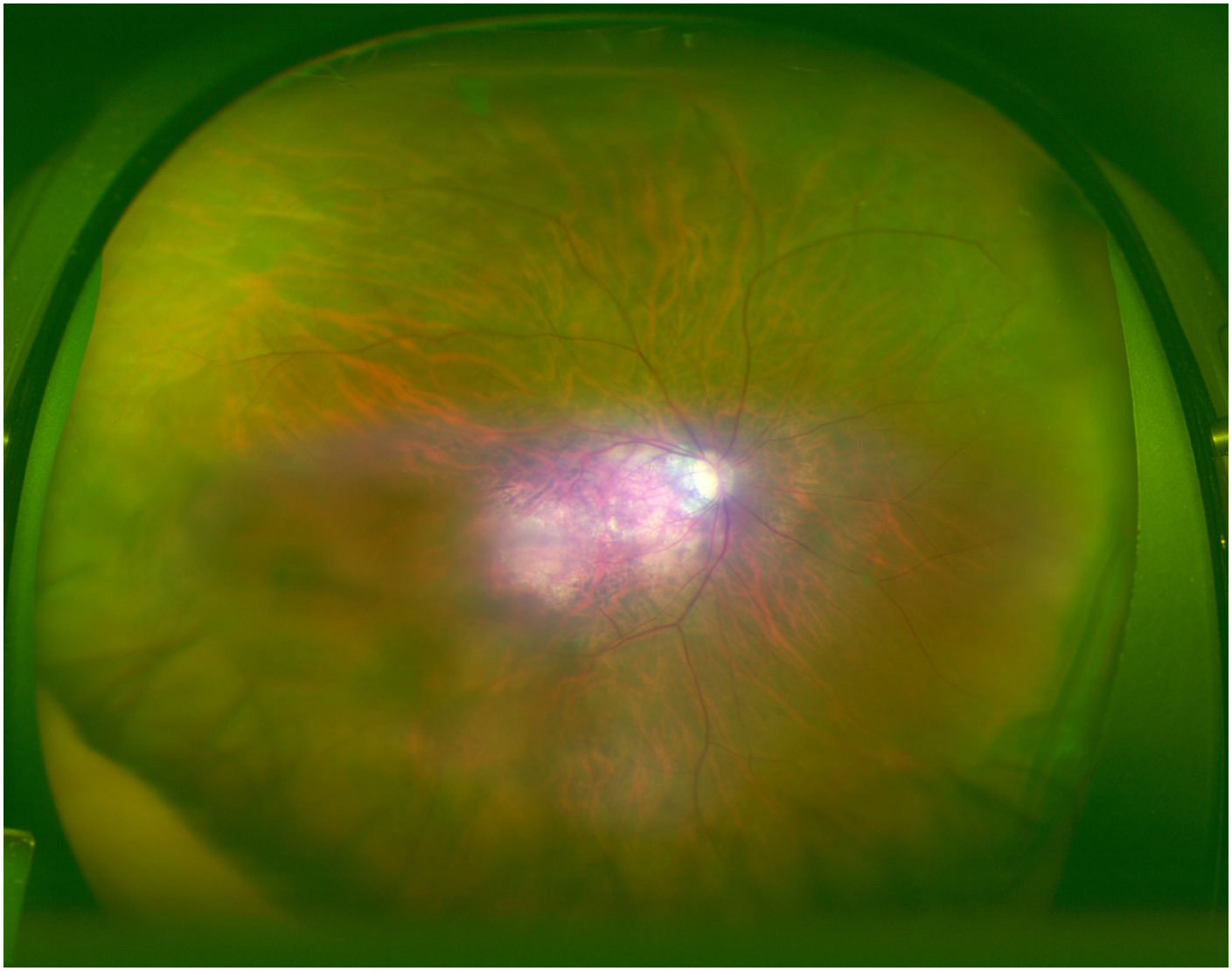

Table 1(on next page)

Demographic data.

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

	Macular hole images	Normal images	p-value	
n	195	715		
Age	66.9 ± 7.6 (20~85)	67.3 ± 12.2 (11～94)	0.5726	Student's *t*-test
Sex (female)	117 (60%)	390 (54.6%)	0.1933	Fisher's exact test
Eye (left)	102 (52.3%)	361 (50.5%)	0.6865	Fisher's exact test

1

2

Table 2(on next page)

The results of CNN model and overall ophthalmologist.

The convolutional neural network model, discrimination test of the macular holes data and the normal data, ophthalmologist, accuracy, sensitivity, specificity, and measurement time.

1

	CNN model	Overall Ophthalmologist
Accuracy	100%	80.6±5.9%
Specificity	100%	95.2±4.3%
sensitivity	100%	69.5±15.7%
measurement time (sec)	32.80±7.36	838.00±199.16

2