**Deep-learning classifier with ultrawide-field fundus ophthalmoscopy for detecting branch retinal vein occlusion**

Daisuke Nagasato¹, Hitoshi Tabuchi¹, Hideharu Ohsugi¹, Hiroki Masumoto¹, Hiroki Enno², Naofumi Ishitobi³, Tomoaki Sonobe¹, Masahiro Kameoka¹, Masanori Niki³, Yoshinori Mitamura³

¹Department of Ophthalmology, Saneikai Tsukazaki Hospital, Himeji 6711227, Japan
²Rist Inc., Tokyo 1530063, Japan
³Department of Ophthalmology, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima 7708503, Japan

**Correspondence to:** Daisuke Nagasato. Department of Ophthalmology, Saneikai Tsukazaki Hospital, 68-1 Waku, Aboshi-ku, Himeji City, Hyogo 671-1227, Japan. d.nagasato@tsukazaki-eye.net

**Received:** 2018-08-18    **Accepted:** 2018-12-06

**Abstract**

- **AIM:** To investigate and compare the efficacy of two machine-learning technologies with deep-learning (DL) and support vector machine (SVM) for the detection of branch retinal vein occlusion (BRVO) using ultrawide-field fundus images.
- **METHODS:** This study included 237 images from 236 patients with BRVO with a mean±standard deviation of age 66.3±10.6y and 229 images from 176 non-BRVO healthy subjects with a mean age of 64.9±9.4y. Training was conducted using a deep convolutional neural network using ultrawide-field fundus images to construct the DL model. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and area under the curve (AUC) were calculated to compare the diagnostic abilities of the DL and SVM models.
- **RESULTS:** For the DL model, the sensitivity, specificity, PPV, NPV and AUC for diagnosing BRVO was 94.0% (95%CI: 93.8%-98.8%), 97.0% (95%CI: 89.7%-96.4%), 96.5% (95%CI: 94.3%-98.7%), 93.2% (95%CI: 90.5%-96.0%) and 0.976 (95%CI: 0.960-0.993), respectively. In contrast, for the SVM model, these values were 80.5% (95%CI: 77.8%-87.9%), 84.3% (95%CI: 75.8%-86.1%), 83.5% (95%CI: 78.4%-88.6%), 75.2% (95%CI: 72.1%-78.3%) and 0.857 (95%CI: 0.811-0.903), respectively. The DL model outperformed the SVM model in all the aforementioned parameters (P<0.001).

- **CONCLUSION:** These results indicate that the combination of the DL model and ultrawide-field fundus ophthalmoscopy may distinguish between healthy and BRVO eyes with a high level of accuracy. The proposed combination may be used for automatically diagnosing BRVO in patients residing in remote areas lacking access to an ophthalmic medical center.

**KEYWORDS:** automatic diagnosis; branch retinal vein occlusion; deep learning; machine-learning technology; ultrawide-field fundus ophthalmoscopy

**DOI:**10.18240/ijo.2019.01.15


**INTRODUCTION**

Branch retinal vein occlusion (BRVO) is a relatively common retinal vascular disorder that causes retinal hemorrhage and macular edema (ME), eventually leading to visual impairment[1-2]. ME resulting from BRVO is associated with poor visual outcomes[3-4]. It has been proposed that a delay in the initiation of ME treatment resulting from BRVO affects functional improvement and hinders improvement in visual acuity[5]. For BRVO, it is important to initiate treatment with antivascular endothelial growth factor agents at an early stage[5-6]. Treatment of patients at a vitreoretinal center shortly after the onset of BRVO is essential for the preservation of visual function. However, establishing vitreoretinal centers that provide such advanced ophthalmological treatments is impractical considering the associated costs burdening social security schemes of numerous nations worldwide[7]. Recently, remarkable progress has been achieved in the development of medical equipment, such as the ultrawide-field scanning laser ophthalmoscope (Optos 200Tx; Optos PLC, Dunfermline, United Kingdom) (Figure 1). The Optos system noninvasively provides wide-field fundus images without using a mydriatic agent and is used for the diagnosis, monitoring,
and treatment of various retinal and choroidal disorders. If there is no risk of elevation of intraocular pressure due to pupil block after mydriasis, the examiner who is not permitted to administer treatment can acquire the images safely. This is ideal for telemedicine applications in areas where there is no ophthalmologist.

In recent years, image processing approaches using two machine-learning algorithms, namely the deep-learning (DL) and support vector machine (SVM) models, have attracted attention because of their extremely high classification performance. Several studies of their application in medical imaging have been conducted. In ophthalmology, the application of an image processing technology using DL to obtain medical images has been previously reported. However, to our knowledge, there have been no studies investigating the automatic diagnosis of BRVO through machine-learning technology using images produced by the Optos system. The aim of this study was to assess the ability of DL and SVM to detect BRVO using Optos images.

SUBJECTS AND METHODS

Ethical Approval  This study was conducted in compliance with the principles of the Declaration of Helsinki and was approved by the Ethics Committees of Tsukazaki Hospital and Tokushima University Hospital. Written informed consents were obtained from all subjects for publication of this study and accompanying images.

Data Set  Optos image data of patients with BRVO and those without fundus diseases were extracted from the clinical database of the ophthalmology departments of Tsukazaki Hospital and Tokushima University Hospital. These images were reviewed by a retinal specialist for the presence of acute BRVO and registered in an analytical database. Of the 466 fundus images selected, 237 belonged to BRVO patients, while 229 belonged to non-BRVO healthy subjects.

In this study, we used K-fold cross validation. This method has been described in detail previously. In brief, the image data were divided into K groups. Subsequently, (K-1) groups were used as training data and one group was used as validation data. This process was repeated K times until each of the K groups became a validation data set. The number of groups (K) was calculated using Sturges’ formula: $K = 1 + \log_2 N$. Sturges’ formula is used to decide the number of classes in the histogram. In this study, $1 + \log_2 237 \approx 8.89$ at BRVO, and $1 + \log_2 229 \approx 8.84$ at non-BRVO. So we categorized these two data into nine groups each.

The images of the training data set were augmented through adjustment of brightness, gamma correction, histogram equalization, noise addition, and inversion. Image augmentation increased the amount of learning data by 18-fold. The deep convolutional neural network (DNN) model was created and trained using the data from the preprocessed images.

Deep-learning Model and its Training  The DNN model called a visual geometry group-16 (VGG-16) used in the present study is shown in Figure 2. This type of DNN is known to automatically learn local features of images and generate a classification model. The DNN used ImageNet parameters: the weights of blocks 1-4 were fixed, while the weights of block 5 and the fully connected layers were adjusted.
Automatic diagnosis of retinal disease

Achieve high performance even with less data. Fine tuning was used to increase the learning speed and to avoid overfitting. Masks were used for the first fully connected layer. To avoid overfitting, masking was performed with a probability of 25% for the generalization performance, and dropout processing was done so that the model learns to be robust against the target image acquired in previous layers.

Feature vectors of target images acquired in previous layers and the softmax function for binary classification. To improve generalization performance, dropout processing was done so that masking was performed with a probability of 25% for the first fully connected layer. Fine tuning was used to increase the learning speed and achieve high performance even with less data\[25-26\]. We used parameters from ImageNet: blocks 1 to 4 were fixed, while block 5 and the fully connected layers were trained.

The weights of block 5 and the fully connected layers were updated using the optimization momentum stochastic gradient descent algorithm (learning coefficient=0.0005, inertial term=0.9)\[27-28\]. Of the 40 DL models obtained in 40 learning cycles, the one with the highest rate of correct answers for the test data was selected as the DL model to be evaluated in this study. Keras (https://keras.io/ja/) was run on TensorFlow (https://www.tensorflow.org/), which is written in the Python programming language, to build and evaluate the model.

Support Vector Machine We used the soft-margin SVM implemented in the scikit-learn library using the radial basis function (RBF) kernel\[29\]. We decreased the dimensionality of the images to 60 dimensions. This was the number of dimensions achieved by the highest rate of correct answers for the test data (10-70 dimensions in steps of 10 were tested). Optimal values for the cost parameter C of the SVM and parameter $\gamma$ of the RBF were determined through grid search using tritracation cross validation. The combination with the highest average rate of correct answers was selected. The parameter values for C (1, 10, 100, and 1000) and $\gamma$ (0.0001, 0.001, 0.01, 0.1, and 1) were tested. The final learning model was generated using the optimal parameter values $C=10$ and $\gamma=0.001$.

Validation In each split, we calculated the answers of these models for the validation data and we collected the answers for 466 fundus images (237 BRVO images and 229 normal images) from those answers. In each split, training data and validation data are completely separated.

Outcome Receiver operating characteristic (ROC) curves were constructed based on the abilities of the DL and SVM models to distinguish between BRVO and non-BRVO images. These curves were evaluated using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and area under the curve (AUC).

Statistical Analysis For comparison of Student’s $t$-test was used to compare the age between patients, while Fisher’s exact test was used to compare the sex ratio and the ratio of right to left eye images. The 95%CI of the AUC was obtained as follows. Images judged to exceed a threshold were defined as positive for BRVO, and an ROC curve was constructed. We produced nine models and nine corresponding ROC curves. For AUC, a 95%CI was obtained by assuming a normal distribution and using the average and standard deviation of the nine ROC curves. For sensitivity and specificity, the optimal cutoff values, i.e. the points closest to the point at which both sensitivity and specificity are 100% in each ROC curve, were used\[20\]. The sensitivities and specificities determined at those cutoff values were used. The ROC curve was calculated using scikit-learn, and the CIs for sensitivity and specificity were determined using scipy. The paired $t$-test was used to compare the AUCs of the DL and SVM models.

Statistical Analysis We used the soft-margin SVM implemented in the scikit-learn library using the radial basis function (RBF) kernel\[29\]. We decreased the dimensionality of the images to 60 dimensions. This was the number of dimensions achieved by the highest rate of correct answers for the test data (10-70 dimensions in steps of 10 were tested). Optimal values for the cost parameter C of the SVM and parameter $\gamma$ of the RBF were determined through grid search using tritracation cross validation. The combination with the highest average rate of correct answers was selected. The parameter values for C (1, 10, 100, and 1000) and $\gamma$ (0.0001, 0.001, 0.01, 0.1, and 1) were tested. The final learning model was generated using the optimal parameter values $C=10$ and $\gamma=0.001$.

Validation In each split, we calculated the answers of these models for the validation data and we collected the answers for 466 fundus images (237 BRVO images and 229 normal images) from those answers. In each split, training data and validation data are completely separated.

Outcome Receiver operating characteristic (ROC) curves were constructed based on the abilities of the DL and SVM models to distinguish between BRVO and non-BRVO images. These curves were evaluated using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and area under the curve (AUC).

Statistical Analysis For comparison of Student’s $t$-test was used to compare the age between patients, while Fisher’s exact test was used to compare the sex ratio and the ratio of right to left eye images. The 95%CI of the AUC was obtained as follows. Images judged to exceed a threshold were defined as positive for BRVO, and an ROC curve was constructed. We produced nine models and nine corresponding ROC curves. For AUC, a 95%CI was obtained by assuming a normal distribution and using the average and standard deviation of the nine ROC curves. For sensitivity and specificity, the optimal cutoff values, i.e. the points closest to the point at which both sensitivity and specificity are 100% in each ROC curve, were used\[20\]. The sensitivities and specificities determined at those cutoff values were used. The ROC curve was calculated using scikit-learn, and the CIs for sensitivity and specificity were determined using scipy. The paired $t$-test was used to compare the AUCs of the DL and SVM models.

Heat Map Images were created by overlaying heatmaps of the DNN focus site on the corresponding BRVO and non-BRVO images. A heatmap of the DNN image focus sites was created and classified using gradient-weighted class activation mapping\[30\]. The target layer is as the third convolution layer in block 3. The ReLU is represented as backprop_modifier. This process was performed using Python Keras-vis (https://raghakot.github.io/keras-vis/).

RESULTS In total, 237 BRVO images from 236 patients (mean age: 66.3±10.6y; 123 males and 113 females; 119 left fundus images and 118 right fundus images) and 229 non-BRVO images from 176 patients (mean age: 64.9±9.4y; 92 males and 84 females; 110 left fundus images and 119 right fundus images) were analyzed. There were no significant differences observed between the two groups in terms of age, sex ratio, or the ratio of right to left eye images. There was a significantly higher rate of hypertension, diabetes and arterial occlusive disorders in the BRVO group than in the non-BRVO group (Table 1).

<table>
<thead>
<tr>
<th>Table 1 Patient demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
</tr>
<tr>
<td>No. of images (patients)</td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>Sex, female (%)</td>
</tr>
<tr>
<td>Left fundus (%)</td>
</tr>
<tr>
<td>Hypertension (%)</td>
</tr>
<tr>
<td>Diabetes (%)</td>
</tr>
<tr>
<td>Arterial occlusive disorders (%)</td>
</tr>
</tbody>
</table>

BRVO: Branch retinal vein occlusion.
The sensitivity of the DL model for the diagnosis of BRVO was 94.0% (95% CI: 93.8%-98.8%), the specificity was 97.0% (95% CI: 93.8%-98.7%), the PPV was 96.5% (95% CI: 94.3%-98.7%), the NPV was 93.2% (95% CI: 90.5%-96.0%) and the AUC was 0.976 (95% CI: 0.960-0.993). In contrast, for the SVM model, these values were 80.5% (95% CI: 77.8%-87.9%), 84.3% (95% CI: 75.8%-86.1%), 83.5% (95% CI: 78.4%-88.6%), 75.2% (95% CI: 72.1%-78.3%) and 0.857 (95% CI: 0.811-0.903). In the ROC curves, the AUC of the DL model was significantly better than that of the SVM model ($P<0.001$; Figure 3).

An image with the corresponding heat map superimposed was produced by the DNN, and the focused coordinate axes in the image were indicated. A representative image is presented in Figure 4. In the image without BRVO, the focal points accumulated around the optic disc. On the other hand, in the image with BRVO, the focal points accumulated around the optic disc and retinal hemorrhages. It is suggested that the DNN may distinguish a BRVO eye from a healthy eye by focusing on the retinal hemorrhages. Blue color was used to indicate the strength of the DNN attention. In the Optos images, the intensity of the color increased on the area of retinal hemorrhages and accumulation was noted at the focus points.

**DISCUSSION**

In this study, the DL model detecting BRVO through the use of Optos fundus images showed higher sensitivity, specificity, PPV, NPV and AUC than the SVM model. DL is known to automatically recognize the local feature values of images and generate classification models\cite{21,25,28,31}. Additionally, DL includes several layers for the identification of local features of complicated differences, which can subsequently be combined\cite{28}. Wang et al\cite{32} reported that the performance of the DL model in the classification of mediastinal lymph node metastases of non-small-cell lung cancer using positron emission tomography/computed tomography images was not significantly different from that of the best standard methods, including SVM and human doctors. In the field of ophthalmology, we previously showed that the DL model using the DNN achieved a better AUC than the SVM model for the detection of rhegmatogenous retinal detachment using ultrawide-field fundus images\cite{15}. The present study confirmed that the performance of the DL model using the DNN was better than that of the SVM model. This result indicates the possibility of early detection of BRVO through combination of Optos fundus images with DL. Our results demonstrated a classification performance of the DL model that was close to that based on the judgment of an ophthalmologist. At the heat map, the DNN focused around the optic disc in the non-BRVO Optos fundus images and around the optic disc and retinal hemorrhages in the BRVO Optos fundus images. This finding suggests that the proposed DNN model may be useful in diagnosing BRVO by identifying suspected retinal hemorrhages caused by BRVO.

The interpretation of all the acquired Optos fundus photographic images by an ophthalmologist is impractical and costly. However, screening for BRVO may be conducted by non-physician personnel in a nonmydriatic and noninvasive manner using the proposed approach. The combination of the DL model and ultrawide-field fundus ophthalmoscopy is a cost-effective option for the screening and diagnosis of large numbers of patients. This approach may be particularly useful for the diagnosis of BRVO in areas with a shortage or lack of ophthalmic care.

Maa et al\cite{33} reported that tele-ophthalmology has the potential to improve operational efficiency, reduce cost, and significantly improve access to care. In areas with a
Automatic diagnosis of retinal disease

shortage of ophthalmological care, the availability of an Optos system offers noninvasive ultrawide-field fundus imaging without requiring the use of a mydriatic agent and avoids the occurrence of complications. The DL technology is able to perform accurate diagnoses of BRVO at a high rate using Optos images. Patients diagnosed with BRVO using this method can immediately consult a retinal specialist and receive the necessary advanced treatment at an ophthalmic medical center. This approach will permit early intervention in BRVO patients residing in medically underserved areas. Moreover, this tele-ophthalmologic technology using Optos may preserve good visual function in BRVO patients residing in areas with inadequate ophthalmic care worldwide.

The following limitations of this study must be acknowledged. Firstly, this study compared only images of health retinas and retinas with BRVO. It did not include images of retinas with other fundus diseases. For an expanded application of this model in the clinical setting, an investigation of other retinal diseases is necessary. Secondly, the analytical ability of Optos is compromised in cases with disorders reducing the clarity of the eye, such as dense cataracts or severe vitreous hemorrhage. Hence, such images were not included in this study. Finally, it is necessary to conduct studies with larger sample sizes and include research on images of other fundus diseases for a more comprehensive evaluation of the performance and versatility of the DL model.

In conclusion, the combination of the DL model and ultrawide-field fundus ophthalmoscopy may distinguish between healthy and BRVO eyes with a high level of accuracy.

ACKNOWLEDGEMENTS

Authors’ contributions: Nagasato D wrote the main manuscript text. Tabuchi H, Ohsugi H and Enno H designed the research. Tabuchi H and Mitamura Y conducted the research. Masumoto H performed the DL methods, the SVM methods, and the statistical analysis. Ishitobi N, Sonobe T, Kameoka M and Niki M collected the data. All authors reviewed the manuscript.

Conflicts of Interest: Nagasato D, None; Tabuchi H, None; Ohsugi H, None; Masumoto H, None; Enno H, None; Ishitobi N, None; Sonobe T, None; Kameoka M, None; Niki M, None; Mitamura Y, None.

REFERENCES


