Predicting Branch Retinal Vein Occlusion Development: A Multimodal Deep Learning Using Pre-onset Retinal Vascular Hemisection Images

Eun Young Choi

The Graduate School
Yonsei University
Department of Medicine

Predicting Branch Retinal Vein Occlusion Development: A Multimodal Deep Learning Using Pre-onset Retinal Vascular Hemisection Images

A Dissertation Submitted
to the Department of Medicine
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Medical Science

Eun Young Choi

June 2024

This certifies that the Dissertation of Eun Young Choi is approved.

Thesis Supervisor	Min Kim
Thesis Committee Member	Junwon Lee
Thesis Committee Member	Hae-Jeong Park
Thesis Committee Member	Song Ji Hun
Thesis Committee Member	Yae Won Park

The Graduate School
Yonsei University
June 2024

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor, Prof. Min Kim, who provided insightful guidance and feedback throughout this project. I would also like to express my appreciation to Dr. Tae Keun Yoo for his practical assistance with deep learning analysis. Additionally, I extend my heartfelt thanks to my husband, Prof. Jinyoung Yeo, for his unwavering support and encouragement throughout the entire process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Predicting Branch Retinal Vein Occlusion Development: A Multimodal Deep Learning Using Pre-onset Retinal Vascular Hemisection Images

Purpose: Branch retinal vein occlusion (BRVO) is a major cause of visual impairment in working-age individuals. In addition to systemic vascular risk factors, retinal vascular features are considered predisposing factors for BRVO. However, predicting BRVO occurrence based solely on retinal vascular features remains elusive. We aimed to develop a deep learning model for BRVO prediction using pre-onset, metadata-matched pairs of fundus hemisection images.

Methods: We retrospectively collected fundus images obtained before disease onset, including 27 paired hemisection images from affected and 81 from unaffected eyes, to use as fundamental learning data. A U-net model was constructed to segment the retinal optic disc and blood vessels (BVs). Fundus and BV segmented images were divided into upper and lower halves based on the optic disc center and labeled according to BRVO occurrence. A BV-enhanced multimodal convolutional neural network model for BRVO prediction was developed using 108 fundus and 108 BV hemisection images.

Results: The BV-enhanced multimodal model achieved an area under the receiver operating characteristic curve (AUC) of 0.76 (95% confidence interval [CI], 0.66–0.83) and accuracy of 68.5% (95% CI, 58.9–77.1%), outperforming the unimodal models based solely on fundus (AUC=0.68) or BV images (AUC=0.72). The multimodal model mainly focused on the retinal vascular arcade, particularly arteriovenous crossing sites.

Conclusions: Incorporating U-net BV segmentation enabled the development of a deep learning model for predicting BRVO occurrence using a metadata-matched dataset. The BV-enhanced multimodal model could predict BRVO occurrence from fundus images. Compiling a learning dataset of fundus images obtained before BRVO onset through a multicenter study is expected to improve the model's predictive performance.

Key words : arteriovenous crossing, blood vessel segmentation, branch retinal vein occlusion, convolutional neural networks, deep learning, retinal vascular features

I. INTRODUCTION

Retinal vein occlusion (RVO) is the second most common retinal vascular disorder after diabetic retinopathy, representing an important cause of visual impairment among the working-age population.[1] Branch RVO (BRVO), characterized by the occlusion of one of the main branches of the central retinal vein, is the most prevalent form, with a reported global prevalence of 4.4 per 1000 individuals and 5-year incidence rate of 0.6% in the United States.[2,3] The severity of visual impairment due to BRVO largely depends on the location, extent, and duration of the occlusion. Ischemic BRVO is associated with an average visual acuity of 20/50, while nonischemic BRVO reduces visual acuity to 20/60.4 Acute BRVO, characterized by intraretinal hemorrhage and cystoid macular edema, can cause visual field defects, which may spontaneously resolve in some cases.[5] However, chronic macular edema, ischemic maculopathy, and neovascularization, which are secondary complications of BRVO, can lead to severe and irreversible visual acuity reduction.5 Despite the use of intravitreal anti-vascular endothelial growth factor agents, periocular or intravitreal corticosteroids, and laser photocoagulation for managing these secondary complications, vision recovery remains limited even with aggressive local treatments.[6]

Old age has been identified as a prominent risk factor for BRVO.[7] According to prior research, the development of BRVO is more closely associated with various cardiovascular high-risk conditions than with ocular factors.[1] Ocular risk factors for the development of BRVO include retinal vascular abnormalities, particularly arteriovenous crossing patterns and nicking.[8,9] Arteriovenous crossing refers to the intersecting of retinal arteries and veins within the retina, while arteriovenous nicking indicates compression of a retinal vein by a retinal artery. These vascular abnormalities can cause venous stenosis, increasing the risk for impaired venous flow and thrombus formation, potentially contributing to BRVO development.[10]

Despite the identification of various systemic and ophthalmic risk factors associated with BRVO, accurately predicting the risk for BRVO occurrence remains challenging. Furthermore, predicting the location of BRVO onset based solely on fundus images is currently deemed nearly impossible. In most cases, BRVO is confined to the arteriovenous crossing of either the superior or inferior vascular arcade of a single eye, with a limited probability of subsequently affecting the other eye of approximately 10%.[11,12] This tendency of BRVO occurring only in specific regions of one eye may be due to the intraindividual differences in retinal vascular features, particularly in individuals with specific systemic risk factors. Therefore, we hypothesized that employing deep learning techniques to extract relevant retinal vascular features from fundus images obtained before BRVO onset could be used to predict the occurrence of BRVO.

Accordingly, in this study, we aimed to develop a deep learning model that can predict the occurrence of BRVO through analysis of fundus images. When collecting image data to construct a deep leaning model, some image data, such as retinal vascular features, may be overlooked due to various factors, including changes caused by age or systemic diseases like hypertension or diabetes mellitus. Therefore, we used a metadata-matched hemisection dataset for model construction. Deep learning using fundus data with matching metadata will be able to predict BRVO development by focusing on structural factors of retinal images. Identifying potential predictors of BRVO in fundus images can enable early prediction and promote targeted

interventions before disease onset. Consequently, the findings of this research are expected to have significant implications for advancing medical knowledge related to BRVO and improving personalized preventive care.

I I. MATERIALS AND METHODS

2.1. Study Design and Population

The study was approved by the Institutional Review Board of Severance Hospital (approval no. 4-2023-1366) and conformed with the tenets of the Declaration of Helsinki. The requirement for informed consent was waived due to the retrospective design and use of anonymized data.

A retrospective analysis of imaging data was conducted using fundus images obtained before disease onset in patients diagnosed with unilateral BRVO at the Ophthalmology Departments of Severance Eye Hospital and Gangnam Severance Hospital between 2005 and 2023. BRVO is a condition in which the eye presents a retinal hemorrhage or other biomicroscopic evidence of RVO, such as telangiectatic capillary bed or dilated venous system in one quadrant or less of the retina drained by the affected vein.[13]

The study population included adults aged 20 years and older with healthy fundus examination results before their BRVO diagnosis. Due to difficulties in identifying vascular structures on the nasal side of the optic disc in fundus photography centered on the macula, we selectively included only BRVO cases occurring in the temporal region, the primary vascular arcade area that feeds the macula. Cases with the presence of other retinal diseases or significant media opacity that obscures retinal vascular features were excluded from the study population. The research team directly collected and anonymized fundus images from the institutional database through the Severance Clinical Research Analysis Portal according to the inclusion criteria, ensuring the absence of any personally identifiable information. At least two board-certified ophthalmologists confirmed that the fundus images obtained before and after BRVO onset belonged to the same person by comparing their blood vessel (BV) and optic nerve morphologies. Low-quality images of the fundus in which BVs could not be identified were excluded.

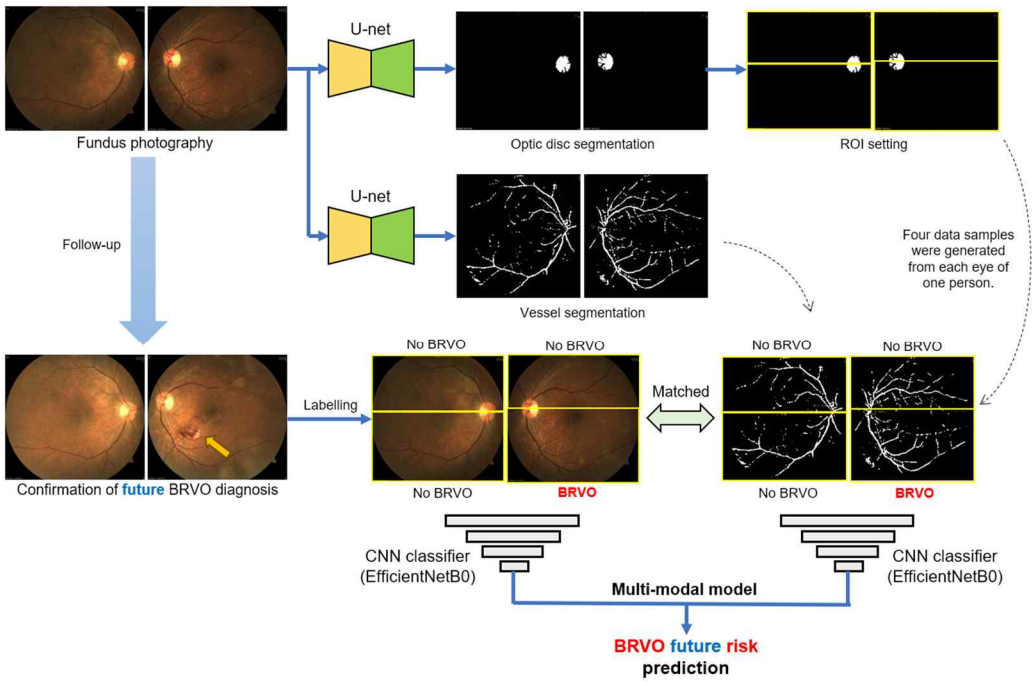
We obtained bilateral fundus images of each eligible patient and divided them into upper and lower halves (hemisections) based on a horizontal line passing through the optic disc center. Retinal vascular features, including arteriovenous crossing pattern (arterial or venous overcrossing) and arteriovenous distance, were analyzed using the integrated calipers (Heidelberg Eye Explore software) in each hemisection image independently. The intersection angles at arteriovenous crossings were independently measured in each hemisection image using ImageJ software (National Institutes of Health, Bethesda, MD, USA).

2.2 Data Processing

2.2.1. Deep learning model for image segmentation

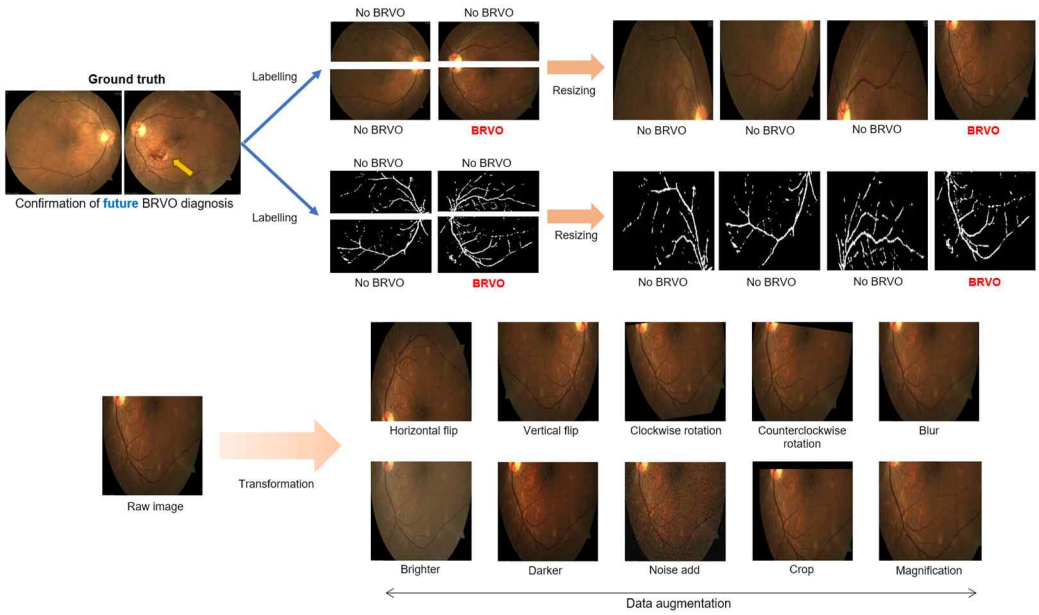
The overall data processing flow is shown in <Fig 1>. A development set was prepared based on fundus image data collected from patients with BRVO before disease onset. PyTorch (Python

library) and Google Colaboratory development environments were employed for deep learning design and training.



<Fig 1> Data processing in a deep learning prediction approach using fundus photography data from those who showed branch retinal vein occlusion (BRVO) during a follow-up period.

We adopted the U-net deep learning segmentation model to extract optic disc and retinal BV images and to calculate the position of the optic disc center.[14] U-net stands out as a fundamental deep learning model, demonstrating robust image segmentation capabilities, particularly in the medical domain. The U-net architecture and related tasks are illustrated in <Fig 2>.



<Fig 2> Top and bottom separation process of retinal images based on optic nerve location and data augmentation.

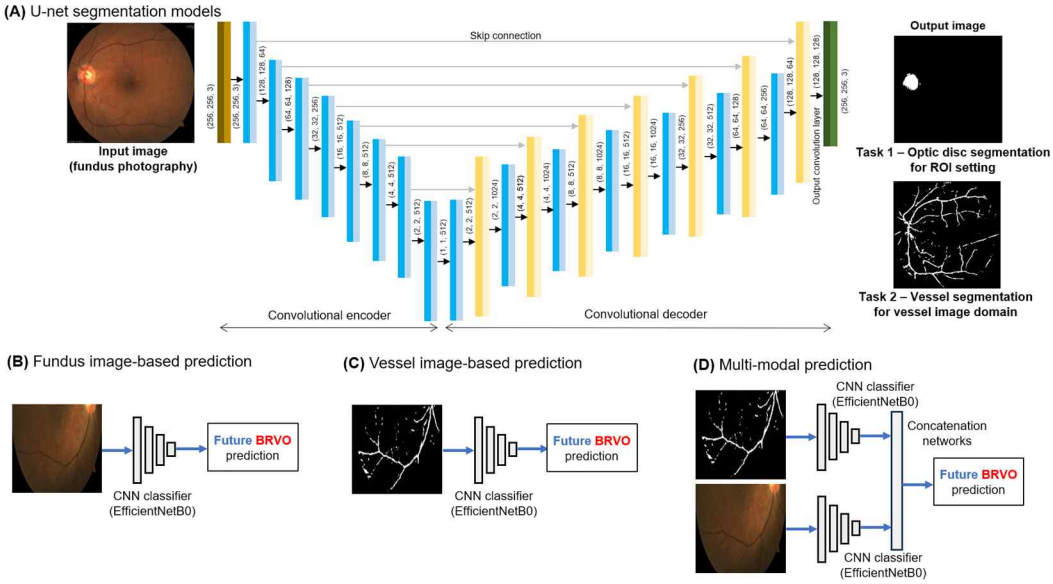
As depicted in Appendix 1, we developed a segmentation model for each task by leveraging well-known external datasets. BV segmentation was trained using the DRIVE and FIVES datasets.[15,16] For optic disc segmentation training, we used the REFUGE challenge dataset.[17] Models trained on these datasets were subsequently applied to the present dataset, resulting in the generation of BV and optic disc segmented images.

Retinal and BV segmented images were divided into upper and lower halves based on the y-coordinate of the pixel center of the optic disc segmentation result (Appendix 3). Each hemisection image was resized into a square with dimensions of 224×224 pixels for deep learning input and used as one training datum, and labeled according to BRVO occurrence, as verified by follow-up data. This enabled us to use the conventional deep learning architecture, which has a state-of-the-art classification performance, without modifying its structure. Subsequently, convolutional neural network (CNN)-based deep learning classifiers were trained to predict the risk for BRVO occurrence using the labeled hemisection images.

2.2.2 Deep learning model for predicting BRVO occurrence

Two CNN models were constructed, one utilizing fundus and one utilizing BV hemisection images (<Fig 3>). Each unimodal model was built using EfficientNetB0 as the backbone, the most recently employed CNN architecture in this field (Appendix 4).[18,19] Each image datum assigned to

the training set underwent a data augmentation process to mitigate overfitting and enhance the generalization ability of the prediction model. Various data augmentation methods were applied, including left and right flipping, random rotation within a +/- 30-degree range, blurring, brightening, darkening, random noise insertion, horizontal movement cropping, and random enlargement and reduction within a +/- 10% range.



<Fig 3> Deep learning model architectures used in this study. (A) U-Net for optic disc segmentation and vessel segmentation. (B) Fundus photography image-based future branch retinal vein occlusion (BRVO) prediction. (C) Vessel segmentation image-based prediction. (D) Multimodal approach using both fundus photography and vessel segmentation images.

Subsequently, an additional BV-enhanced multimodal model combining the two image domains was constructed using a concatenation network following the protocols from previous studies.[20,21] For concatenating the CNN outputs, the last layers of the two trained EfficientNetB0 models for each image domain were substituted with modified fully connected network layers (512, 128, and 52 nodes with dropout) and two SoftMax functions for the two classes (future BRVO occurrence or not). These functions set the output of the prediction score to a range of zero to one, corresponding to the prediction probability (Appendix 5). The concatenation network was further trained using pairs of fundus and matched BV hemisection images.

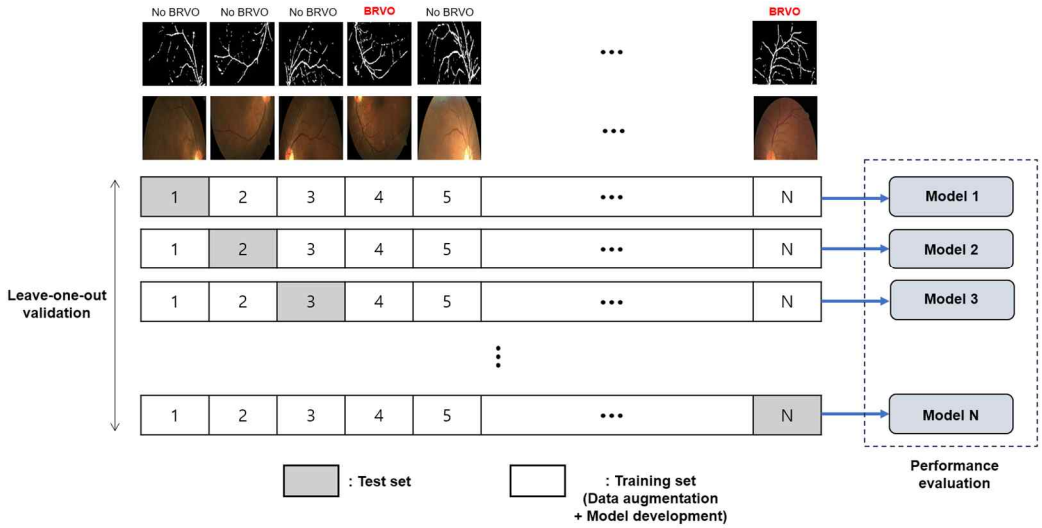
The EfficientNetB0 models were initially pretrained on ImageNet data and then imported into the workspace. The EfficientNet series trained on ImageNet are models that have achieved state-of-the-art performance in image classification problems, and the weights of the models are all publicly available, allowing them to be used in various domain-specific tasks. The models were stably trained on a small dataset through a transfer learning approach.[22] All network training procedures were

optimized using stochastic gradient descent (SGD) with a momentum algorithm (SGD learning rate = 0.0001) and a mini-batch size of 20 over 100 epochs, which are standard fine-tuning parameters for transfer learning. Using the gradient-weighted class activation map (Grad-CAM) technique for each trained EfficientNetB0 model, attention maps were generated from the last layers of the SoftMax and activation convolutional layers. This Grad-CAM visualization indicates whether the deep learning model was properly trained with a focus on features related to the occurrence of BRVO. Model development and validation were conducted using PyTorch.

2.3. Statistical Analysis

The segmentation models were assessed using the intersection-over-union (IoU) and Dice coefficient. The performance of the BRVO prediction models was evaluated using binary classification metrics, including the area under the receiver operating characteristics curve (AUC), accuracy, sensitivity, and specificity. Youden's index was adopted to select the diagnostic threshold, assigning equal weights to sensitivity and specificity.

As illustrated in <Fig 4>, the performance evaluation of the BRVO prediction models was conducted using the leave-one-out cross-validation method. This approach involves creating a total of N models (equals the number of samples), excluding only one sample when building each model, calculating test set performance with the excluded sample, and averaging the performance of the N trained models.[23] Leave-one-out cross-validation is a technique that avoids random sampling, providing a better estimate of model performance when the dataset is small. This technique is effective for evaluate the generalization ability of deep learning models and avoids overfitting and underfitting when small datasets are used.[24] The softmax outputs from each leave-one-out cross-validation were collected to plot the receiver operating characteristic curve.



<Fig 4> Schematic diagram of data processing using leave-one-out algorithm for studying top-bottom separation and multimodal analysis of retinal images.

III. RESULTS

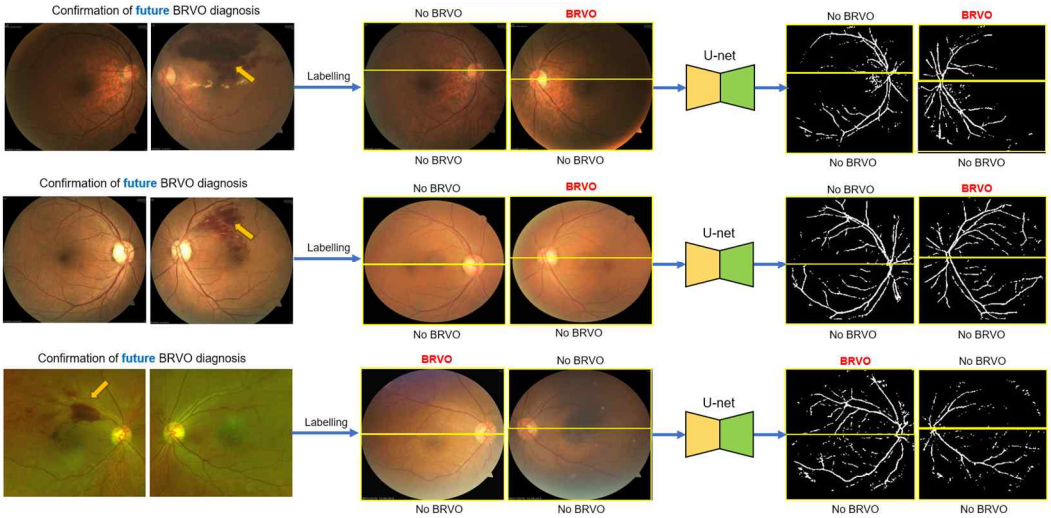
Among 2,673 patients diagnosed with unilateral BRVO whose images were initially collected, only 31 had available unaffected fundus images obtained before BRVO diagnosis. Four patients were excluded from the training dataset because the retinal BVs in the macula could not be accurately extracted due to wide-field imaging or poor image quality. Hence, the final training dataset consisted of fundus image data from 27 patients with BRVO. The demographic and clinical characteristics of these patients are summarized in Appendix 2. Each fundus image was segmented into upper and lower halves, referred to as hemisections. Three pairs of unaffected hemisections were derived for each BRVO-affected hemisection. Consequently, the training dataset comprised 27 hemisections affected by BRVO in the future and 81 paired hemisections that did not develop BRVO. The matched pair design with metadata control ensured that both BRVO-affected and unaffected data were collected from the same patients. This approach eliminated the influence of demographic and systemic clinical factors on the deep learning model analysis. Therefore, it was not feasible to construct a prediction model, such as logistic regression, using metadata.

<Table 1> presents the performance of the segmentation models. For the BV segmentation task, the performance of the U-net model trained on data from DRIVE and FIVES was evaluated using the respective validation sets (Appendix 6). The trained model exhibited IoU and Dice coefficient values of 0.805 and of 0.921 in the DRIVE test set, and 0.824 and 0.947 in the FIVES validation set. For the optic nerve segmentation task, the U-net model trained on the REFUGE set demonstrated an IoU value of 0.933 and Dice coefficient of 0.962 in the validation set. Using these trained segmentation models, we successfully generated segmented upper and lower hemisection fundus and BV image datasets for BRVO prediction. Examples of retinal BV segmentation using the trained U-net model are shown in <Fig 5>.

<Table 1> Segmentation performance using U-Net for retinal vessel segmentation and optic disc segmentation.

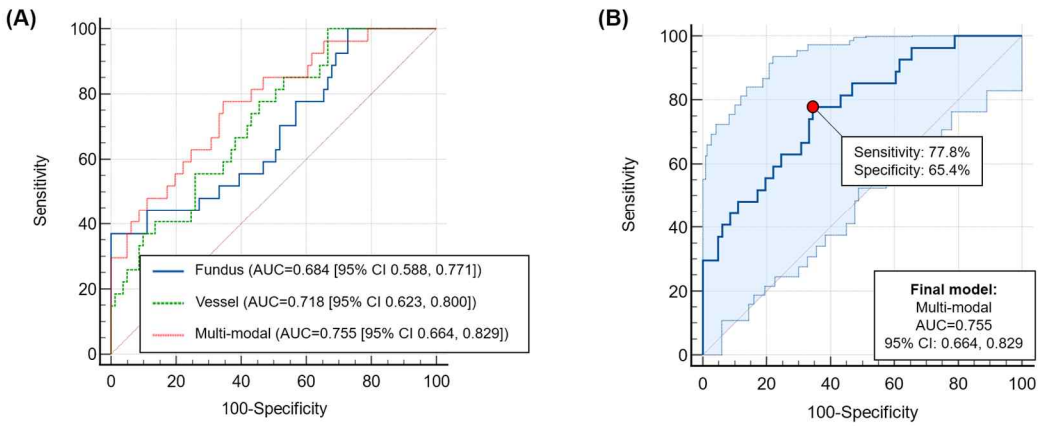
Model	Validation set	IoU	Dice coefficient
U-net for vessel segmentation	DRIVE test set [15]	0.805	0.921
	FIVES dataset (randomly selected 20% for validation) [16]	0.824	0.947
U-net for optic disc segmentation	REFUGE test set [17]	0.933	0.962

IoU, intersection-over-union.



<Fig 5> Example of retinal blood vessel segmentation results applying the trained U-net model.

The receiver operating characteristic curves, drawn using the prediction results extracted from the leave-one-out cross-validation, are presented in <Fig 6>. The multimodal model achieved an AUC value of 0.755 (95% confidence interval [CI], 0.664–0.829), outperforming the two unimodal models based on fundus and BV hemisection images, which achieved AUC values of 0.684 (95% CI, 0.588–0.771) and 0.718 (95% CI, 0.623–0.800), respectively. All three models demonstrated significant predictive power for BRVO ($P < 0.05$), with 95% CI intervals of the AUC values of random sampling greater than 0.5.



<Fig 6> Leave-one-out validation performance of the developed models to predict BRVO.

development. (A) Receiver operating characteristics (ROC) curves comparison. (B) The ROC curve of the multimodal deep learning prediction model.

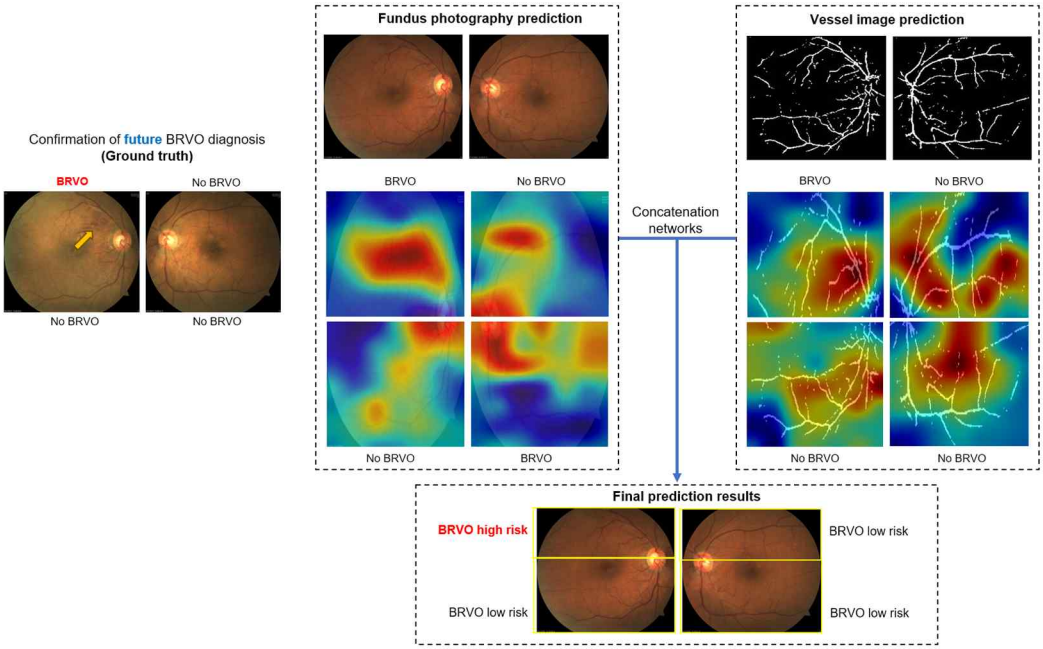
The details of the performance evaluation of the prediction models are presented in <Table 2>. The BV-enhanced multimodal CNN model demonstrated the highest accuracy of 68.5% (95% CI, 58.9–77.1%), with a sensitivity of 77.8% (95% CI, 57.4–91.4%) and specificity of 65.4% (95% CI, 54.0–75.7%). The accuracy, sensitivity, and specificity of the unimodal CNN models were 45.4% (95% CI, 35.8–55.2%), 100.0% (95% CI, 87.2–100.0%), and 27.2% (95% CI, 17.9–38.2%) for the fundus image-based model, and 50.0% (95% CI, 40.2–59.8%), 100.0% (95% CI, 87.2–100.0%), and 33.3% (95% CI, 23.2–44.7%) for the BV image-based model, respectively.

<Table 2> Leave-one-out validation performance of the developed models to predict future branch retinal vein occlusion.

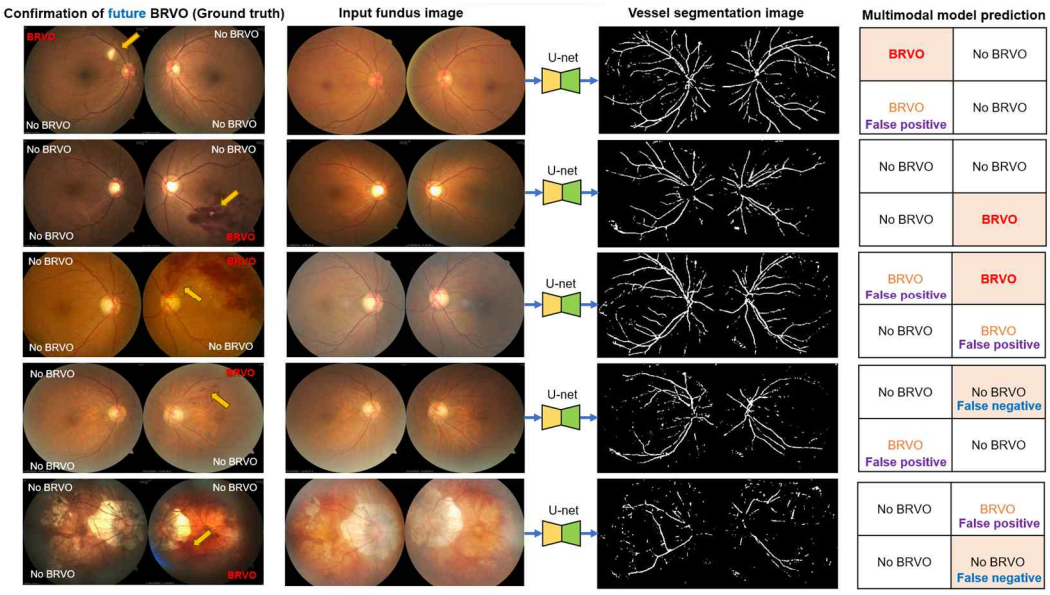
Deep learning models	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
CNN based on fundus images	0.684 (0.588–0.771)	45.4 (35.8–55.2)	100.0 (87.2–100.0)	27.2 (17.9–38.2)
CNN based on vessel images	0.718 (0.623–0.800)	50.0 (40.2–59.8)	100.0 (87.2–100.0)	33.3 (23.2–44.7)
Multimodal CNN	0.755 (0.664–0.829)	68.5 (58.9–77.1)	77.8 (57.4–91.4)	65.4 (54.0–75.7)

AUC, area under the curve; CI, confidence interval; CNN, convolutional neural network

<Fig 7> presents model attention maps based on the Grad-CAM technique in the prediction of BRVO. Each hemisection of the fundus and BV segmented images was analyzed to visualize where the model focused. In 74.1% of true-positive BRVO cases among hemisection BV images, the decision was made by focusing on the retinal vascular arcades, including the arteriovenous crossing site (Grad-CAM signal more than 0.75). However, in the remaining cases, no specific pattern was identified. As depicted in <Fig 8>, arteries and veins crossed at small angles in the retinal vascular arcade, and a coarse BV segmentation area at the crossing site was frequently associated with BRVO occurrence. To further analyze the deep learning model's focus on the arteriovenous crossing area, we examined all arteriovenous crossing angles in each hemisection. The average arteriovenous crossing angle was 21.9 ± 13.3° in the area where BRVO occurred, 33.2 ± 20.8° in the counter hemisection, and 31.3 ± 11.6° in the contralateral eye. The angle in the area where BRVO occurred was significantly smaller than that in the counter hemisection (P=0.0391) and the average of the contralateral eye (P=0.0072).



<Fig 7> Attention maps based on the Grad-CAM technique to predict future branch retinal vein occlusion (BRVO) using fundus photograph and vessel segmentation images.



<Fig 8> Input fundus image, retinal blood vessel segmentation image, and prediction result of the final multimodal deep learning model. True positive, false positive, false negative, and true negative cases were sampled.

IV. DISCUSSION

In this study, we developed a multimodal deep learning model enhanced by retinal BV information for predicting BRVO using fundus images collected from patients who later developed BRVO. We could predict future BRVO occurrence in pre-onset healthy eyes by analyzing retinal vascular features alone, using paired fundus hemisection datasets. This paired hemisection dataset was perfectly matched in terms of metadata, enabling the development of deep learning models that focus exclusively on the retinal vasculature structure. The multimodal model predicted BRVO development with fairly high accuracy, primarily by focusing on retinal vascular features. Our BV-enhanced multimodal model has two main features in design. First, we utilized U-net to segment retinal BVs; this allows the CNN model to focus on the running structure of arterial and venous vessels, known to be associated with BRVO occurrence.8 Second, the unimodal models predicted BRVO from either the fundus or BV segmented images; on the contrary, the BV-enhanced multimodal model used a combination of both predictions. The clinical significance of this study lies in the creation of a model capable of predicting BRVO occurrence based on normal fundus images. Because we used fundus images obtained before BRVO onset as training data, the segmentation models were developed from a dataset adjusted for age and cardiovascular disease risk. Although the accuracy of the unimodal models in predicting BRVO occurrence was not high, the BV-enhanced multimodal model demonstrated a significant improvement in prediction performance.

The strength of this study is that we used fundus images of patients with confirmed diagnosis of BRVO obtained before disease onset and leveraged deep learning to extract vascular structural features—known risk factors for BRVO—for prediction. Studying the morphological features of retinal BVs in eyes with developed BRVO is challenging due to the tortuous dilation, capillary nonperfusion, collateral formation, and obscuration by retinal hemorrhage and edema.[1] Previous cohort studies have prioritized cardiovascular risk factors over the morphological features of retinal BVs in relation to BRVO development.[25–27] However, with advancements in technologies like optical coherence tomography (OCT) and OCT angiography, morphological characteristics such as vascular crossing or stenosis were found to be significant risk factors for BRVO.[28,29] Moreover, this study is strengthened by the composition of the training dataset, which consisted of paired fundus hemisection images from the same patient, presenting both BRVO-affected and unaffected areas. This allowed us to focus on the structural aspects of retinal BVs within data that was perfectly matched for metadata. The significance of this approach is that it facilitated metadata-matched deep learning prediction. Considering that BRVO is significantly influenced by age and systemic chronic diseases, the design of this dataset addressed critical challenges commonly faced by deep learning models analyzing fundus images.

According to the Grad-CAM analysis and prediction results, the deep learning model in this study predicted BRVO occurrence mainly by focusing on arteriovenous crossings at acute angles. This pattern is particularly prominent in the CNN prediction results based on BV segmented hemisection images. BRVO development has been associated with venous compression at sites of

arterial or venous overcrossing at arteriovenous intersections.[9,28,30] Previous retrospective studies analyzing fundus images obtained before BRVO onset have identified specific BV characteristics in cases where BRVO occurred, including arteriovenous crossings, smaller arteriovenous crossing angles, double vessel crossings with venous stasis, and crossings near retinal vein bifurcation.[8,10] Quantitative analysis of each of these morphological BV features in individual fundus images is challenging, but these features could be learned from images using the BV-enhanced multimodal deep learning algorithm for predicting BRVO.

When evaluating the performance of the multimodal model for BRVO prediction, we observed instances of false-positive and false-negative predictions (<Fig 8>). Predicting the occurrence of a disease, usually influenced by various factors, using limited biomarkers is a challenging task. As our goal was to predict BRVO occurrence solely through deep learning extraction of vascular features from fundus images, other demographic and clinical factors were not considered. Generally, incorporating correlated covariates as learning data can enhance the prediction accuracy of the model. Nevertheless, in this study, the number of patients with fundus examination completed before BRVO onset was rather small, resulting in a small learning dataset; thus, various covariates could not be considered.[31] The diminished predictive power of the model is also attributed to the low quality of fundus images. As depicted in <Fig 8>, incomplete retinal BV segmentation occurred in blurry or dark areas of the images. Moreover, accurate segmentation of retinal BVs was hindered by interference from visible choroidal vessels, even in eyes with myopic retinal changes.

In this study, the BRVO prediction performance significantly improved with a multimodal method enhanced by BV segmentation. The approach of dividing and extracting necessary image information using U-net and then adopting a multimodal strategy instead of directly analyzing the original image using CNN has demonstrated great efficacy in previous studies on medical image analysis.[32] With deep learning models using OCT images, higher accuracy in predicting retinal diagnosis was achieved by segmenting the normal retinal layer and lesion area.[33] Even in glaucoma diagnosis models using fundus images, the diagnostic performance could be enhanced through a multimodal approach involving an optic nerve segmentation model.[34] The method of obtaining overall higher performance on limited datasets by combining the image pattern recognition ability of CNN with the structural segmentation ability of U-net can be extended to various diagnostic areas in the future.[35]

The primary limitation of this study lies in its foundation on a small development dataset for deep learning-based research. Developing models with limited datasets often raises concerns regarding potential performance limitations and the risk of overfitting. To mitigate these issues, our approach utilized bilateral fundus divided into upper and lower halves, forming a metadata-matched paired dataset. This method enabled pairing each BRVO-affected hemisection with three unaffected hemisections, thereby quadrupling the dataset while maintaining control over metadata other than the images. Data augmentation was also leveraged, applying ten transformations to all input images to maximize both the quantity and diversity of the training data. Despite these efforts on the dataset, the overall size of the training data remained relatively small for the demands of deep learning. To address these constraints, the study employed several strategies throughout the model development process. A multimodal approach was implemented to incorporate retinal blood vessel features as crucial supplementary information alongside fundus images. The use of EfficientNetB0, pre-trained on ImageNet, facilitated stable learning from the small dataset. Furthermore, transfer learning was applied in the segmentation of retinal blood vessels, employing a proven U-net segmentation model trained on an external open dataset to generate accurate retinal blood vessel images from our limited

number of fundus images.

Given the rarity of pre-onset fundus photographs for BRVO, it is noteworthy that this study was able to compile a significant dataset spanning 19 years of medical records from two institutions. This effort marks an important step towards developing a predictive model for BRVO, laying the groundwork for future research. With the potential to gather more extensive data from multicenter studies or health examination databases, there is an optimistic outlook for enhancing the accuracy and clinical applicability of the model using the presented multimodal framework. Future research could also explore various methods for integrating the two unimodal structures, as well as developing an advanced segmentation model that can distinguish between retinal arteries and veins, thereby further refining the model's predictive capabilities.

V. CONCLUSION

The present study demonstrated the feasibility of predicting BRVO occurrence through BV-enhanced multimodal deep learning analysis of retinal BV features in fundus images using a paired hemisection dataset. Collecting fundus images before BRVO onset posed challenges but provided valuable learning data. The multimodal architecture was adopted to focus on utilizing structural information of retinal BVs, achieved through BV segmentation based on transfer learning pretrained on the appropriate open datasets. While the prediction performance was not exceptionally high, we verified the potential for predicting BRVO risk using only fundus images obtained under normal conditions. Consistent with existing knowledge, we identified the importance of vascular structures, particularly the arteriovenous crossing area, for BRVO prediction. Incorporating the U-net segmentation model, grounded in medical evidence, contributed technically and academically to the deep learning prediction of a major retinal vascular disease, BRVO. Collecting sufficient data through multicenter studies or population-based surveys, with separate extraction of retinal arteries and veins and consideration of cardiovascular risk factors, holds promise for developing a more accurate BRVO occurrence prediction model.

References

1. Jaulim A, Ahmed B, Khanam T, Chatziralli IP. Branch retinal vein occlusion: Epidemiology, Pathogenesis, Risk Factors, Clinical Features, Diagnosis, and Complications. An Update of the Literature. RETINA. 2013;33(5):901. doi:10.1097/IAE.0b013e3182870c15
2. Laouri M, Chen E, Looman M, Gallagher M. The burden of disease of retinal vein occlusion: review of the literature. Eye. 2011;25(8):981-988. doi:10.1038/eye.2011.92
3. Rogers S, McIntosh RL, Cheung N, Lim L, Wang JJ, Mitchell P, et al. The prevalence of retinal vein occlusion: pooled data from population studies from the United States, Europe, Asia, and Australia. Ophthalmology. 2010;117(2):313-319.e1. doi:10.1016/j.ophtha.2009.07.017
4. Thapa R, Paudyal G, Bernstein PS. Demographic characteristics, patterns and risk factors for retinal vein occlusion in Nepal: a hospital-based case-control study. Clin Exp Ophthalmol. 2010;38(6):583-590. doi:10.1111/j.1442-9071.2010.02295.x
5. Hayreh SS, Rojas P, Podhajsky P, Montague P, Woolson RF. Ocular neovascularization with retinal vascular occlusion-III. Incidence of ocular neovascularization with retinal vein occlusion. Ophthalmology. 1983;90(5):488-506. doi:10.1016/s0161-6420(83)34542-5
6. Chatziralli IP, Jaulim A, Peponis VG, Mitropoulos PG, Moschos MM. Branch retinal vein occlusion: treatment modalities: an update of the literature. Semin Ophthalmol. 2014;29(2):85-107. doi:10.3109/08820538.2013.833271
7. Lee JY, Yoon YH, Kim HK, Yoon HS, Kang SW, Kim JG, et al. Baseline Characteristics and Risk Factors of Retinal Vein Occlusion: A Study by the Korean RVO Study Group. J Korean Med Sci. 2013;28(1):136-144. doi:10.3346/jkms.2013.28.1.136
8. Kawasaki R, Nagano E, Uno M, Okada M, Kawasaki Y, Kitamura A. Retinal Vascular Features Associated with Risk of Branch Retinal Vein Occlusion. Current Eye Research. 2013;38(9):989-993. doi:10.3109/02713683.2013.798420
9. Muraoka Y, Tsujikawa A. Arteriovenous crossing associated with branch retinal vein occlusion. Jpn J Ophthalmol. 2019;63(5):353-364. doi:10.1007/s10384-019-00676-5
10. Noh SY, Lee JH, Jeong WJ. Branch Retinal Vein Occlusion with Arteriovenous Crossing. J Retin. 2023;8(1):36-41. doi:10.21561/jor.2023.8.1.36
11. Rogers SL, McIntosh RL, Lim L, Mitchell P, Cheung N, Kowalski JW, et al. Natural History of Branch Retinal Vein Occlusion: An Evidence-Based Systematic Review. Ophthalmology. 2010;117(6):1094-1101.e5. doi:10.1016/j.ophtha.2010.01.058
12. Zhao J, Sastry SM, Sperduto RD, Chew EY, Remaley NA. Arteriovenous Crossing Patterns in Branch Retinal Vein Occlusion. Ophthalmology. 1993;100(3):423-428. doi:10.1016/S0161-6420(93)31633-7
13. Campochiaro PA, Heier JS, Feiner L, Gray S, Saroj N, Rundle AC, et al. Ranibizumab for Macular Edema following Branch Retinal Vein Occlusion: Six-Month Primary End Point Results of a Phase III Study. Ophthalmology. 2010;117(6):1102-1112.e1. doi:10.1016/j.ophtha.2010.02.021
14. Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nat Methods. 2019;16(1):67-70. doi:10.1038/s41592-018-0261-2
15. Staal J, Abramoff MD, Niemeijer M, Viergever MA, van Ginneken B. Ridge-based vessel segmentation in color images of the retina. IEEE Transactions on Medical Imaging. 2004;23(4):501-509. doi:10.1109/TMI.2004.825627
16. Jin K, Huang X, Zhou J, Li Y, Yan Y, Sun Y, et al. FIVES: A Fundus Image Dataset for

Artificial Intelligence based Vessel Segmentation. Sci Data. 2022;9(1):475. doi:10.1038/s41597-022-01564-3
17. Orlando JI, Fu H, Barbosa Breda J, Keer KV, Bathula DR, Diaz-Pinto A, et al. REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. Medical Image Analysis. 2020;59:101570. doi:10.1016/j.media.2019.101570
18. Marques G, Agarwal D, de la Torre Díez I. Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. Applied Soft Computing. 2020;96:106691. doi:10.1016/j.asoc.2020.106691
19. Son J, Shin JY, Kong ST, Park JH, Kwon GT, Kim HD, et al. An interpretable and interactive deep learning algorithm for a clinically applicable retinal fundus diagnosis system by modelling finding-disease relationship. Sci Rep. 2023;13(1):5934. doi:10.1038/s41598-023-32518-3
20. Yoo TK, Choi JY, Seo JG, Ramasubramanian B, Selvaperumal S, Kim DW. The possibility of the combination of OCT and fundus images for improving the diagnostic accuracy of deep learning for age-related macular degeneration: a preliminary experiment. Med Biol Eng Comput. 2019;57(3):677-687. doi:10.1007/s11517-018-1915-z
21. Vaghefi E, Hill S, Kersten HM, Squirrell D. Multimodal Retinal Image Analysis via Deep Learning for the Diagnosis of Intermediate Dry Age-Related Macular Degeneration: A Feasibility Study. Journal of Ophthalmology. 2020;2020:e7493419. doi:10.1155/2020/7493419
22. Jain R, Yoo TK, Ryu IH, Song J, Kolte N, Nariani A. Deep Transfer Learning for Ethnically Distinct Populations: Prediction of Refractive Error Using Optical Coherence Tomography. Ophthalmol Ther. Published online November 13, 2023. doi:10.1007/s40123-023-00842-6
23. Shao Z, Er MJ. Efficient Leave-One-Out Cross-Validation-based Regularized Extreme Learning Machine. Neurocomputing. 2016;194:260-270. doi:10.1016/j.neucom.2016.02.058
24. Yang C, Zhang X, Bruzzone L, Liu B, Liu D, Ren X, et al. Comprehensive mapping of lunar surface chemistry by adding Chang'e-5 samples with deep learning. Nat Commun. 2023;14:7554.
25. Cugati S, Wang JJ, Rochtchina E, Mitchell P. Ten-Year Incidence of Retinal Vein Occlusion in an Older Population: The Blue Mountains Eye Study. Archives of Ophthalmology. 2006;124(5):726-732. doi:10.1001/archopht.124.5.726
26. Chang YS, Ho CH, Chu CC, Wang JJ, Jan RL. Risk of retinal vein occlusion in patients with diabetes mellitus: A retrospective cohort study. Diabetes Research and Clinical Practice. 2021;171:108607. doi:10.1016/j.diabres.2020.108607
27. Kolar P. Risk factors for central and branch retinal vein occlusion: a meta-analysis of published clinical data. J Ophthalmol. 2014;2014:724780. doi:10.1155/2014/724780
28. Muraoka Y, Tsujikawa A, Murakami T, Ogino K, Kumagai K, Miyamoto K, et al. Morphologic and Functional Changes in Retinal Vessels Associated with Branch Retinal Vein Occlusion. Ophthalmology. 2013;120(1):91-99. doi:10.1016/j.ophtha.2012.06.054
29. Kumagai K, Tsujikawa A, Muraoka Y, Akagi-Kurashige Y, Murakami T, Miyamoto K, et al. Three-dimensional optical coherence tomography evaluation of vascular changes at arteriovenous crossings. Invest Ophthalmol Vis Sci. 2014;55(3):1867-1875. doi:10.1167/iovs.13-13303
30. Kogo T, Muraoka Y, Iida Y, Ooto S, Murakami T, Kadomoto S, et al. Angiographic Risk Features of Branch Retinal Vein Occlusion Onset as Determined by Optical Coherence Tomography

Angiography. Investigative Ophthalmology & Visual Science. 2020;61(2):8. doi:10.1167/iovs.61.2.8
31. Martens MJ, Logan BR. A Unified Approach to Sample Size and Power Determination for Testing Parameters in Generalized Linear and Time-to-Event Regression Models. Stat Med. 2021;40(5):1121-1132. doi:10.1002/sim.8823
32. Ni Y, Xie Z, Zheng D, Yang Y, Wang W. Two-stage multitask U-Net construction for pulmonary nodule segmentation and malignancy risk prediction. Quant Imaging Med Surg. 2022;12(1):292-309. doi:10.21037/qims-21-19
33. De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwellet S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med. 2018;24(9):1342-1350. doi:10.1038/s41591-018-0107-6
34. Li Y, Han Y, Li Z, Zhong Y, Guo Z. A transfer learning-based multimodal neural network combining metadata and multiple medical images for glaucoma type diagnosis. Sci Rep. 2023;13(1):12076. doi:10.1038/s41598-022-27045-6
35. Shuvo MdB, Ahommed R, Reza S, Hashem MMA. CNL-UNet: A novel lightweight deep learning architecture for multimodal biomedical image segmentation with false output suppression. Biomedical Signal Processing and Control. 2021;70:102959. doi:10.1016/j.bspc.2021.102959

APPENDICES

APPENDICES 1. Datasets used to develop model for fundus image segmentation tasks.

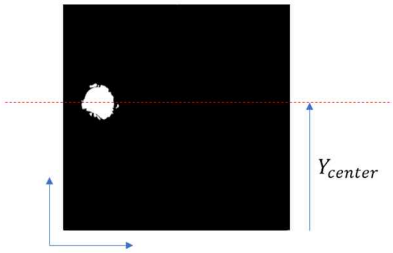
Task	Segmentation training dataset	Summary
Vessel segmentation	DRIVE dataset (Digital Retinal Images for Vessel Extraction) [15]	Fundus photographs and matched binary vessel images Total 40 pairs (20 for training and 20 for test)
	FIVES dataset [16]	Fundus photographs and matched binary vessel images Total 800 pairs
Optic disc segmentation	REFUGE dataset (optic disc detection) [17]	Fundus photographs and matched binary optic disc images Total 400 pairs (320 for training and 80 for test)

APENDICES 2. Demographics and clinical characteristics of the study population (N = 31).

Age, years	69.6 ± 11.3
Gender, female	20 (65)
Underlying comorbidties and past history, N	
Hypertension	18 (58)
Diabetic mellitus	10 (32)
Hyperlipidemia	7 (23)
Cardiovascular diseases	8 (31)
Cerebrovascular diseases	5 (16)
Any cancer history	6 (19)
History of anti-coagulant use	8 (26)
Ever smoking*	3 (9.7)
Ophthalmological information for affected eyes	
Presenting visual acuity, logMAR	0.38 ± 0.41
Spherical equivalent, diopters	-0.80 ± 2.29
Intraocular pressure, mmHg	15.0 ± 5.2
Complications associated with BRVO, yes	
Macular edema	23 (74)
Retinal ischemia	7 (23)
Neovascularization	1 (3.2)
Vitreous hemorrhage	1 (3.23%)
Pattern of BRVO lesions	
Laterality, right	14 (45)
Location	
Superotemporal	18 (58)
Inferotemporal	13 (42)
Mean area, mm^2	19.3 ± 11.8
Distance from the optic disc center of the starting point, μm	2034.9 ± 1217.5
Optical coherence tomography parameters	
Central macular thickness, μm	427.3 ± 166.1
Subfoveal choroidal thickness, μm	204.6 ± 121.8
Fundus fluorescein angiography parameters	
Area of nonperfusion area, disc diameters	0.6 ± 1.6
Ischemic maculopathy, yes	5 (16)

Continuous variables are presented as means and standard deviations, while categorical variables are presented as frequencies and percentages. *Smoking at least 100 cigarettes in his/her lifetime. BRVO, branch retinal vein occlusion; LogMAR, the Logarithm of the Minimum Angle of Resolution.

APENDICES 3. Centroid coordinate extraction to create image hemisections after optic disc segmentation.



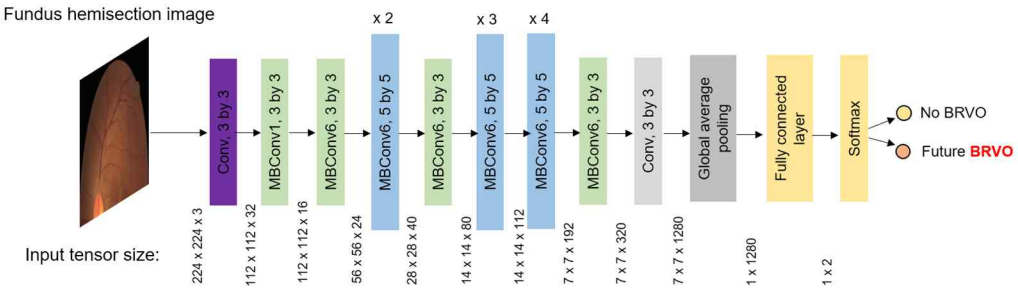
$$Y_{center} = \sum_{i=1}^{N} \frac{I_i \times y_i}{A}$$

y_i = *pixel height*

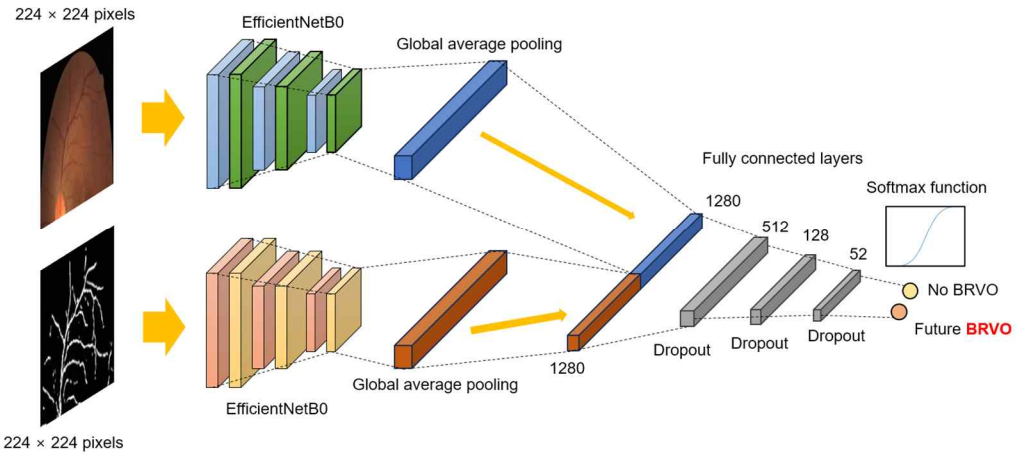
$I_i = 1$, if inside optic disc area
$= 0$, if not

$A = total\ area\ of\ optic\ disc$

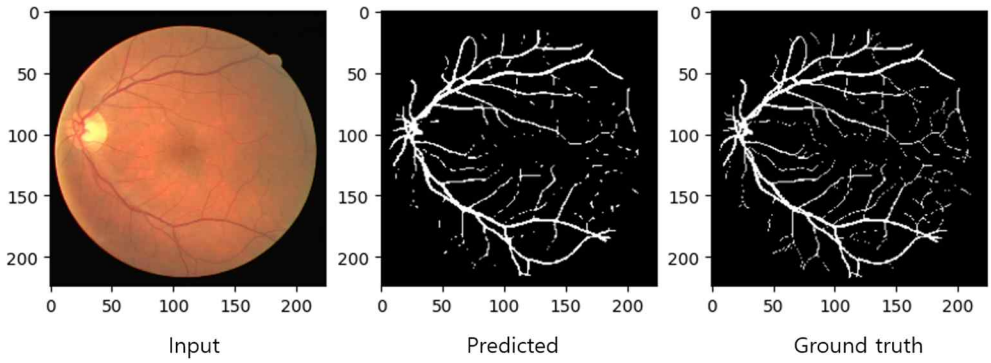
APENDICES 4. Network architecture of EfficientNetB0 for analysis of fundus hemisection images.



APENDICES 5. Network architecture configuration of concatenation network for multimodal analysis for hemisection fundus and blood vessel images.



APENDICES 6. Example of segmented image results from DRIVE dataset using trained U-net model.



Abstract in Korean

발병 전 망막 혈관 이미지의 반단면을 이용해 분지망막정맥폐쇄 발생을 예측하는 멀티모달 딥러닝 모델

목적: 분지망막정맥폐쇄(branch retinal vein occlusion, BRVO)는 근로 연령층에서 시력 손상의 주요 원인이다. 심혈관 질환의 위험 요소와 함께 망막 혈관의 형태적 특성 또한 BRVO 의 위험인자로 알려져 있다. 그러나 망막 혈관 특성만으로 BRVO 발생을 사전에 예측하는 것은 현재로서는 불가능하다. 이에 본 연구는 BRVO 발병 전 안저 사진의 반단면 이미지 쌍을 사용하여 BRVO 를 예측하기 위한 딥러닝 모델 개발을 목표로 한다.

방법: 단안 BRVO 환자의 발병 전 촬영한 양안의 안저 사진을 후향적으로 수집하여 기본 학습 데이터로 활용했다. 전이 학습된 U-net 분할 모델을 사용해 시신경유두와 망막 혈관 이미지를 추출했다. 시신경유두를 중심으로 안저 사진과 혈관 이미지를 상하 반단면으로 나누어, BRVO 발생한 27개의 반단면 이미지와 발생하지 않은 81개의 반단면 이미지를 구분했다. BRVO 예측을 위해, 108개의 안저 이미지 기반 단일모달 모델과 108개의 혈관 반단면 이미지 기반 단일모달 모델을 결합하여 혈관 강화 멀티모달 모델을 개발했다. 모든 모델은 사전학습된 EfficientNetB0를 기반 구조로 사용했다. 모델의 성능을 평가하고 작동 메커니즘을 분석했다.

결과: 혈관 강화 멀티모달 모델은 수용자 작용 특성 곡선(area under the curve, AUC) 0.76(95% 신뢰구간[confidence interval, CI] = 0.66–0.83)과 68.5%의 정확도(95% CI = 58.9–77.1%)를 달성하여, 안저 사진 기반 단일모달 모델(AUC = 0.68) 및 혈관 이미지만을 사용한 모델(AUC = 0.72)보다 우수한 성능을 보였다. Grad-CAM 기술을 활용한 분석 결과 멀티모달 모델은 주로 망막 동정맥 교차 부위에 초점을 맞추며, 예각의 동정맥 교차가 BRVO 발생 예측과 관련이 있음을 보여주었다.

결론: 메타데이터가 일치하는 데이터셋을 구축하고, 전이 학습 기반의 U-net 혈관 분할 및 멀티모달 구조를 통해 망막 혈관에 초점을 맞춘 BRVO 발생 예측 모델을 개발할 수 있었다. 향후 다기관 연구를 통해 BRVO 발생을 보다 정확하게 사전에 예측할 수 있게 되면, 고위험 환자들을 대상으로 한 적극적인 예방 조치가 가능해질 것으로 기대된다.

핵심되는 말 : 동정맥 교차, 혈관 분할, 분지망막정맥폐쇄, 합성곱 신경망, 딥러닝, 망막 혈관 특성