Hindawi
Disease Markers
Volume 2022, Article ID 1730501, 9 pages
https://doi.org/10.1155/2022/1730501

Research Article

Hypermixed Convolutional Neural Network for Retinal Vein Occlusion Classification

Guanghua Zhang,[1,2] **Bin Sun**,[3] **Zhaoxia Zhang**,[3] **Shiyu Wu**,[4] **Guangping Zhuo**,[4] **Huifang Rong**,[1] **Yunfang Liu**,[5] **and Weihua Yang**[6]

[1]*Department of Intelligence and Automation, Taiyuan University, Taiyuan 030000, China*
[2]*Graphics and Imaging Laboratory, University of Girona, Spain*
[3]*Shanxi Eye Hospital, Taiyuan 030002, China*
[4]*Department of Computer, Taiyuan Normal University, Jinzhong 030619, China*
[5]*The First Affiliated Hospital of Huzhou University, Huzhou 313000, China*
[6]*Shenzhen Eye Hospital, Jinan University, Shenzhen 518040, China*

Correspondence should be addressed to Yunfang Liu; panlele0701@139.com and Weihua Yang; benben0606@139.com

Received 30 August 2022; Accepted 30 September 2022; Published 11 November 2022

Academic Editor: Yi Shao

Retinal vein occlusion (RVO) is one of the most common retinal vascular diseases leading to vision loss if not diagnosed and treated in time. RVO can be classified into two types: CRVO (blockage of the main retinal veins) and BRVO (blockage of one of the smaller branch veins). Automated diagnosis of RVO can improve clinical workflow and optimize treatment strategies. However, to the best of our knowledge, there are few reported methods for automated identification of different RVO types. In this study, we propose a new hypermixed convolutional neural network (CNN) model, namely, the VGG-CAM network, that can classify the two types of RVOs based on retinal fundus images and detect lesion areas using an unsupervised learning method. The image data used in this study is collected and labeled by three senior ophthalmologists in Shanxi Eye Hospital, China. The proposed network is validated to accurately classify RVO diseases and detect lesions. It can potentially assist in further investigating the association between RVO and brain vascular diseases and evaluating the optimal treatments for RVO.

1. Introduction

Early changes in the retina are influenced by many factors, such as unfavorable environmental factors, including aging, a high-carbohydrate diet, and a sedentary lifestyle [1], and systemic diseases, including hyperglycemia [2], hyperlipidemia [3], and hypertension [4]. Retinal blood vessels are the only blood vessels available for noninvasive imaging in the human body. The pathological changes in retinal blood vessels occur much earlier than clinical symptomatic lesions. Therefore, retinal images have been widely used to detect early signs of systemic vascular diseases. In recent years, with an increase in the elderly population and the acceleration of the aging society in China, the fundus diseases of the elderly occur more frequently. Impaired vision can significantly impact an older person's quality of life and ability to live independently [5].

Retinal vessels are important structures of our eyes [6], and their detection and analysis are of great significance for the study of ocular diseases. Patients with retinal diseases may exhibit serious complications that cause severe visual impairment owing to a lack of awareness of retinal diseases and limited medical resources [7]. Retinal vein occlusion (RVO) [8] is one of the important eye diseases considered a risk factor for cardiovascular mortality and stroke in aging people [9]. Its typical symptoms include exudate [10], capillary nonperfusion [11], collateral formation [12], microaneurysm [13], sclerosed veins [14], and telangiectatic vessels [15]. Ischemic RVO is usually complicated by macular edema (ME) [16] and retinal and iris neovascularization

[17], resulting in significant visual loss. RVO is classified into central and branch RVO (CRVO and BRVO). CRVO involves superficial or deep retinal hemorrhages (HEs) [18] that are scattered around the vein near the lamina cribrosa [19], and BRVO involves hemorrhages occurring within the occluded venule from the retinal sector to the blood supply sector, which is caused by arterial compression onto veins [20] (see Figure 1). RVO is the second most common retinal vascular eye disease after diabetic retinopathy (DR) [21]. If RVO is not treated in a timely manner, it can lead to serious complications that cause severe visual impairment [22, 23]. So far, the number of patients with RVO has increased [24], but our understanding of its pathogenesis, our ability to modify the final visual outcome, and the availability of treatments to effectively intervene in the progression of the disorder are all relatively limited [25].

Diagnosing ophthalmological diseases through deep learning models [26] has been used broadly in recent years [27]. A convolutional neural network (CNN) is one of the most famous deep learning architectures designed in 1989 [28]. Krizhevsky et al. [29] trained a large CNN architecture with eight layers and millions of parameters using a large ImageNet data set containing 1 million training images. In the field of ophthalmology, Litjens et al. [27] used a CNN to automatically segment macular edema based on OCT images. Google used a CNN network to automate the classification of diabetic retinopathy [30] and obtained the experimental results of 99% referral accuracy by training more than 100,000 data sets. The technology has been approved by the FDA as an official medical product. CNN is good at extracting feature information of different colors, spaces, and edges of images by using convolution modules of different scales and integrating all features into higher-order abstract features of images through continuous nonlinear transformation combinations. High-order abstract features and basic features are used together in the final learning process. CNN is almost a conventional method for medical image analysis, including color fundus images (CFIs) [27]. It is better in configuring spatial information by taking images as input. The achievements of CNN in autodiagnosis on different medical aspects can be found in [27], and it has been proved to surpass humans in some cases. So far, a hierarchical CNN architecture capable of distinguishing between normal CFIs and BRVO CFIs has been proposed by Zhang et al. [31] and developed by Zhao et al. [32]. Over the past several years, many modified and deeper CNN architectures have been proposed, which are not only used in the medical imaging domain but also widely applied in other domains [33].

Our study proposed an advanced model that can classify all normal, CRVO, and BRVO CFIs and detect the visible hemorrhage areas. It will help ophthalmologists realize computer-aided diagnosis in pathological analysis [34] to alleviate their pressure and discover and treat RVO as early as possible [35]. We will illustrate the methods employed in our model in the next chapter, followed by experimental results and final remarks in the third and fourth chapters, respectively.

2. Materials and Methods

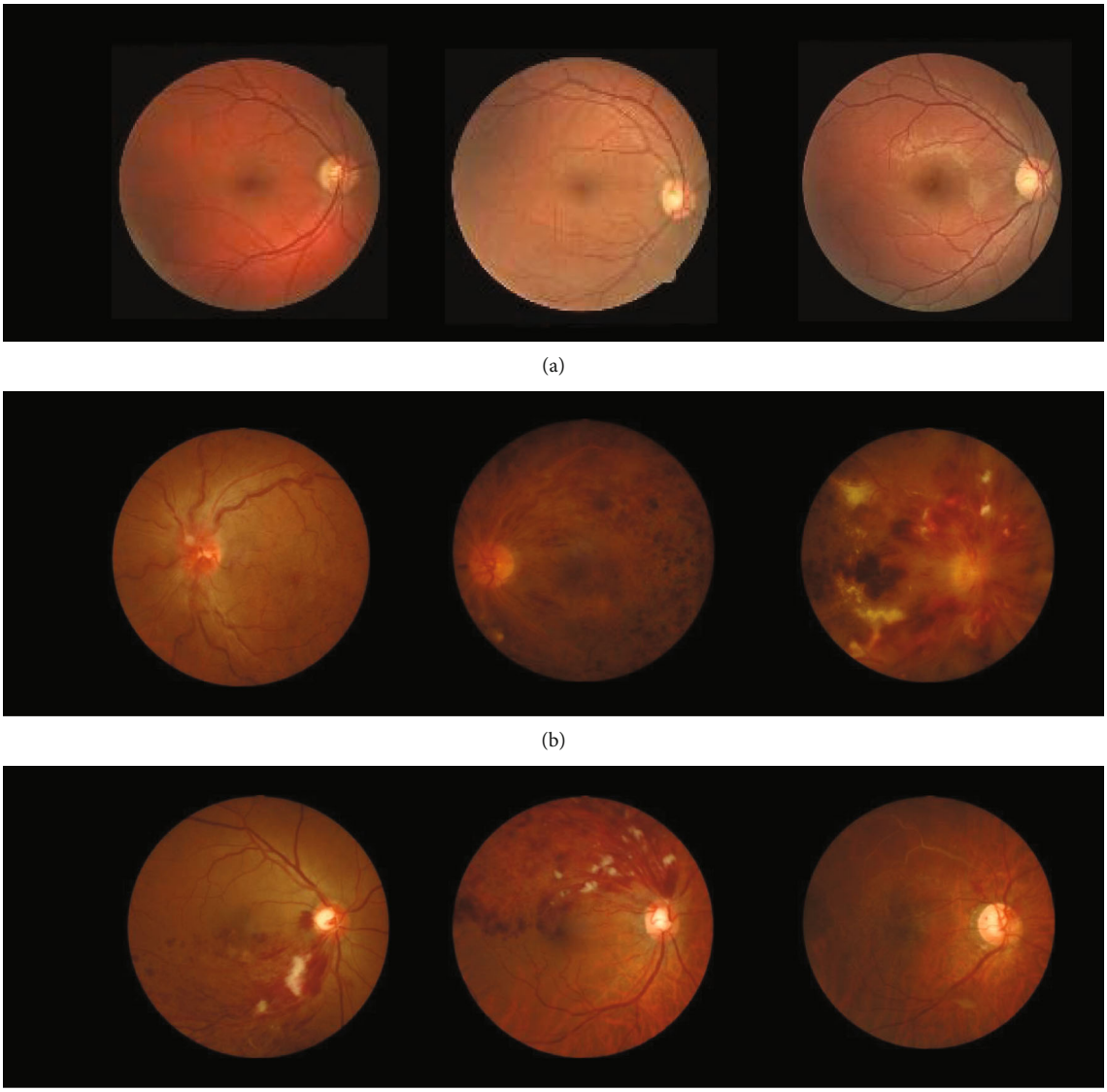
When dealing with medical images, the structural and configuration information between adjacent pixels is a great source for analysis. A CNN that combines convolutional layers, pooling layers, and fully connected layers is more capable of extracting this type of information from 2D or 3D images. Convolutional layers apply a convolution operation by processing images only in receptive fields and adapting the weights gradually during the learning process. Pooling layers are usually followed by convolutional layers to reduce the dimensions of their output. Fully connected layers flatten data from previous layers to one dimension. A simple CNN framework is shown in Figure 2.

In this study, we also use the same data sets to conduct the control experiments on Resnet-34 [37], Inception-V3 [38], and MobileNet [39] models. Resnet [40] is the champion of the ImageNet large-scale visual recognition challenge (ILSVRC) in 2015. Resnet-34 [37] model is mainly composed of residual blocks, through which a deep network can be built and residual learning can be carried out in the feature extraction process. The Inception model is a deep CNN architecture proposed by Szegedy et al. [41] in ILSVRC 2014. The asymmetric multiconvolution kernel structure of the Inception-V3 [38] model performs the splitting operation on the larger convolution. Convolution kernels with different sizes are adopted so that receptive fields of different sizes can exist. The calculation efficiency of model parameters is improved, and the overfitting of the model is reduced. The MobileNet [39] model was proposed by a Google team [42] in 2017 and consists of a series of basic deep separable convolution (DSC) units. The model has a high precision and involves a small number of parameters and calculations.

2.1. Model Architecture. The study introduces a new CNN framework to classify RVO types and detect lesions. It is known as the VGG-CAM network, which utilizes a modified VGG19 network, general average pooling (GAP), class activation mapping (CAM), and CAM attention.

VGG19 [43] is a CNN architecture introduced by Simonyan and Zisserman. They used small receptive fields (3×3 matrix) to detect features from different positions of images and added the number of convolutional layers to increase the reception area for these receptive fields. Our VGG-CAM network reduces the number of fully connected layers in the original VGG19 networks from three to one and replaces them with a GAP layer. In the feature extraction stage, the auxiliary classifier and CAM attention layer are introduced to further enhance the model's activation weight for the lesion area. An additional CAM layer is connected with the GAP layer for lesion detection. SoftMax is applied as an activation function for the final fully connected layer, which predicts probabilities of different classes that the CFI can be labeled. The 24-layer framework of the VGG-CAM network is shown in Figure 3.

In the VGG-CAM network, the GAP layer preserves more information from input images [44], which helps detect lesion areas in the input image. Compared with average pooling, GAP only outputs one parameter from each



(a)

(b)

(c)

Figure 1: (a) Normal retina images, (b) CRVO retina images, and (c) BRVO retina images.

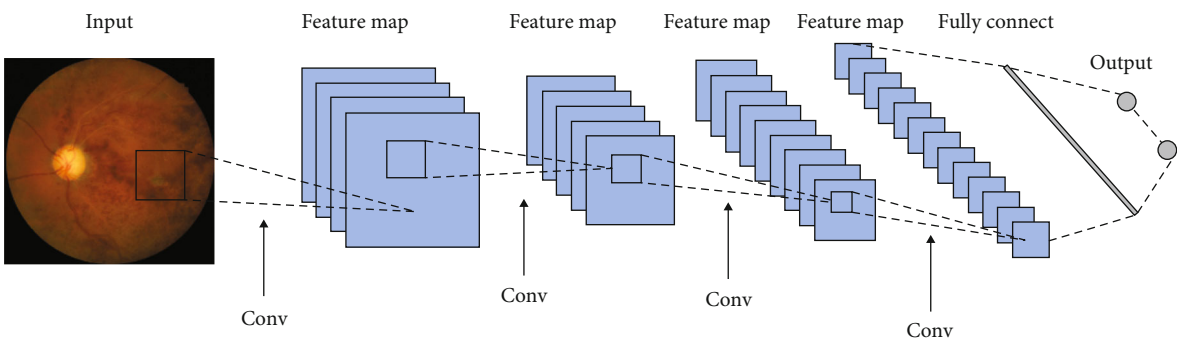


Figure 2: A simple CNN framework containing input, convolutional, pooling (subsampling), and fully connected layers (Heung-II, [36]).

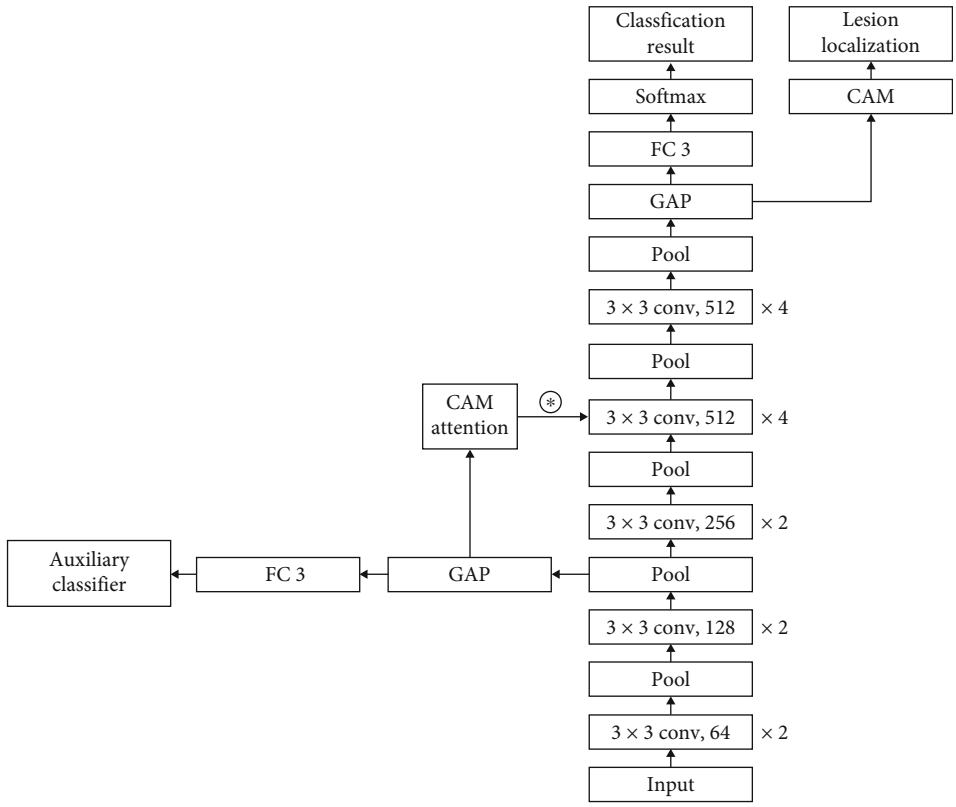


Figure 3: Framework of VGG-CAM model for RVO classification and lesion detection.

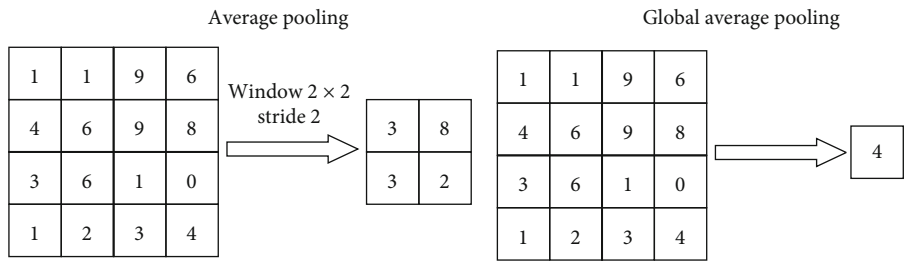


Figure 4: Examples of average pooling and GAP.

receptive field (see Figure 4). Compared with the former, it uses less time to optimize the network. Lin et al. proved that the reduction of parameters in GAP has no effect on the accuracy of final networks [44].

The CAM layer's computation is as follows:

$$\mathrm{CAM} = \sum_{i=0}^{C} \omega_i * F_i, \tag{1}$$

where C represents the number of channels in the feature map of the previous GAP layer. For each feature channel, the CAM layer augments products of weights ω_i from the fully connected layer and feature maps F_i before the GAP layer, as seen in Figure 5.

The CAM layer first uses Equation (1) to compute the class activation image of the original CFIs. It then applies bilinear interpolation to turn the class activation image into the size of the original image, followed by a threshold segmentation to detect the lesion location.

In the feature extraction stage, some useless information often affects the final classification accuracy; further, the extraction of key pathological features is the key to classification. The attention mechanism can guide the model to independently select the lesion area to be noticed. The feature weights are generated by introducing the auxiliary classifier and CAM attention layer after the sixth pooling layer. The feature weight is introduced into the feature map of the 10th layer to improve the model's attention and learning of the lesion area and further improve the final classification effect and the accuracy of lesion detection.

2.2. Image Preprocessing. Input images were mainly preprocessed by contrast limited adaptive histogram equalization

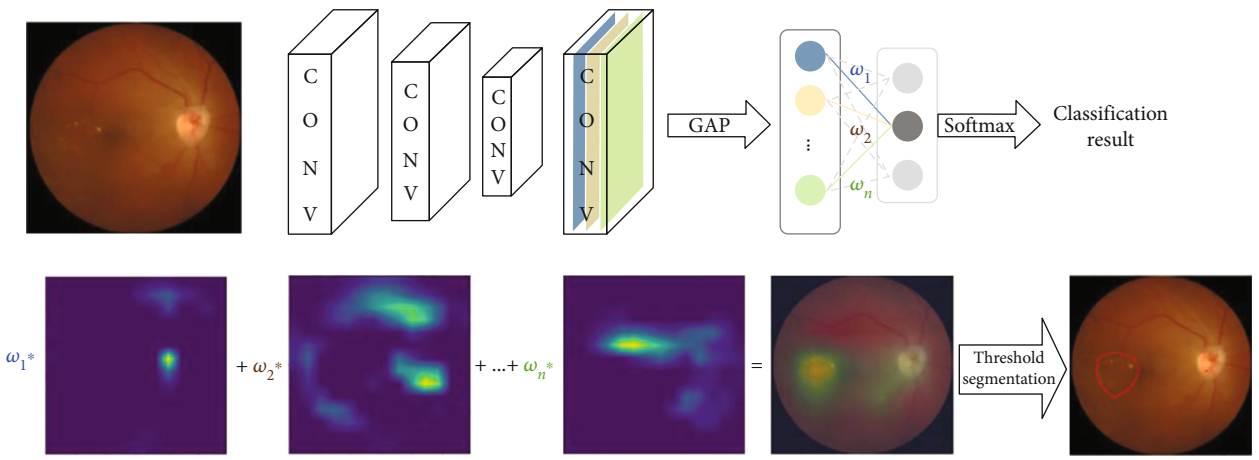


Figure 5: RVO lesion detection by CAM.

(CLAHE) to increase contrast in original images [45]. Flipping, twisting, and zooming were used as well to increase the variety in our image database, which improves the model's ability to recognize various RVO images. Figure 6 presents the differences between original and preprocessed input images.

2.3. Model Initialization

2.3.1. Transfer Learning. Transfer learning [46] applies pretrained weights on a network from another problem as the initial weights for the same network in a different problem. In the problem of image processing, the shallow network of a neural network is mainly responsible for the feature extraction of shallow elements in an image, such as points, edges, and other such elements. Universal pretrained weights can reduce the network's learning time on a different problem [47]. The VGG-CAM model used pretrained weights from ImageNet as initial weights, which was trained by over a million images containing over 1,000 labeled images (https://www.image-net.org/).

2.3.2. Stage-Wise Training. Stage-wise training [48] assigns priority to features of images. It separates the entire learning process into several sublearning processes, and the ability to extract different levels of image features is achieved through different learning processes. It allows information from the images to be processed gradually in the model [49]. In the earliest stage (the first eight layers in our model), the network accessed only a subset of the image, especially its coarse-scale features. Following stage II (8th to 13th layers) and stage III (13th to 18th layers), finer information was extracted from the image, and the feedback was used to evolve the previous stages for a better prediction. Stage III is the only prior for feature learning of the final stage (the fully connected layer).

2.3.3. Environment. Operating system: Ubuntu 18.04 LTS; language: Python 3.6.8, Keras; GPU: GTX1080ti; CPU: Intel i7; Memory: Kingston DDR4 16 G.

3. Results and Discussion

3.1. Evaluation Metrics. In this experiment, the ability to identify unsupervised lesions was tested first, and then, the classification performance of the VGG-CAM model was tested in terms of sensitivity (Se), specificity (Sp), and Kappa. The calculation formulas of each index are shown below.

$$\mathrm{Se} = \frac{\mathrm{TP}}{\mathrm{FN} + \mathrm{TP}}, \tag{2}$$

$$\mathrm{Sp} = \frac{\mathrm{TN}}{\mathrm{FP} + \mathrm{TN}}, \tag{3}$$

$$\mathrm{Kappa} = \frac{P_0 - P_e}{1 - P_e}. \tag{4}$$

In Equations (2) and (3), TP indicates that the positive class is predicted as the positive class number, TN indicates that the negative class is predicted as the negative class number, FN indicates that the negative class is predicted as the positive class number, and FP indicates that the negative class is predicted as the positive class number.

In Equation (4), P_0 represents the sum of the number of samples correctly classified for each class divided by the total number of samples, that is, the overall classification accuracy. Assume that the real number of samples of each class is $a_1, a_2, \cdots, a_n$, respectively, the predicted number of samples of each class is $b_1, b_2, \cdots, b_n$, respectively, and the total number of samples is n; then, P_e is expressed as

$$P_e = \frac{a_1 \times b_1 + a_2 \times b_2 + \cdots + a_n \times b_n}{n \times n}. \tag{5}$$

3.2. Lesion Detection. Figure 7 presents examples of a detected lesion in BRVO and CRVO CFIs. For BRVO CFIs, the VGG-CAM network is sufficiently capable of identifying exudate, sclerosed veins, and hemorrhages. However, the network only highlights parts of a hemorrhage when the hemorrhage area is large. As for CRVO that has hemorrhage spreading all over the retina, the VGG-CAM network only

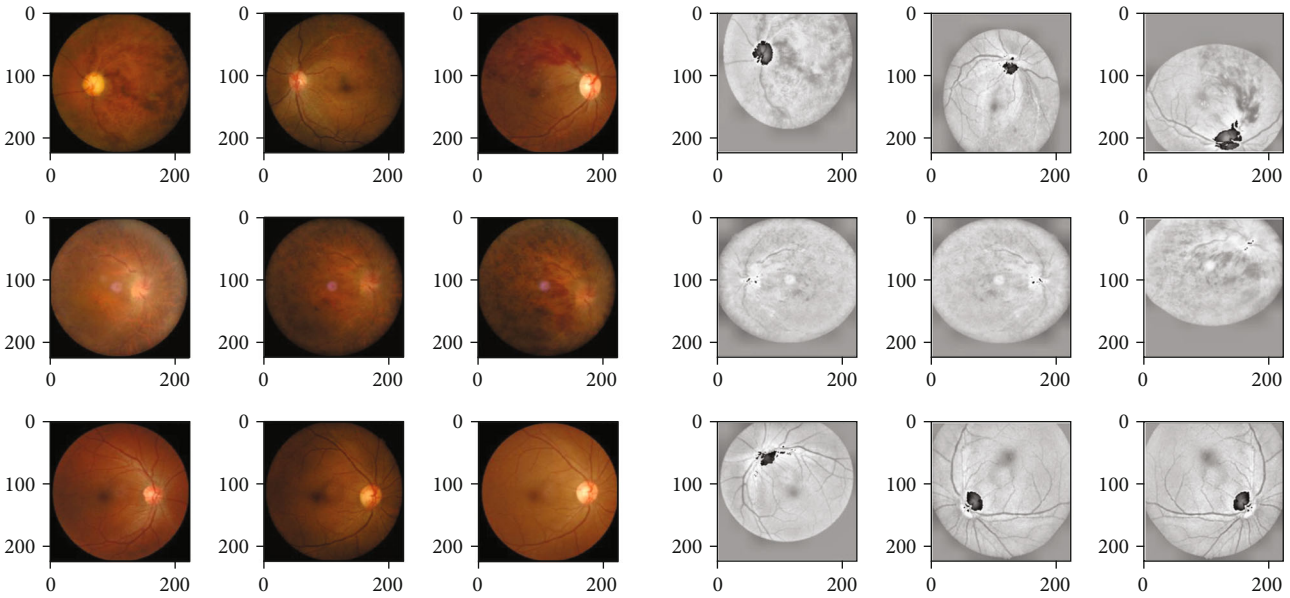


Figure 6: Input images before and after preprocessing.

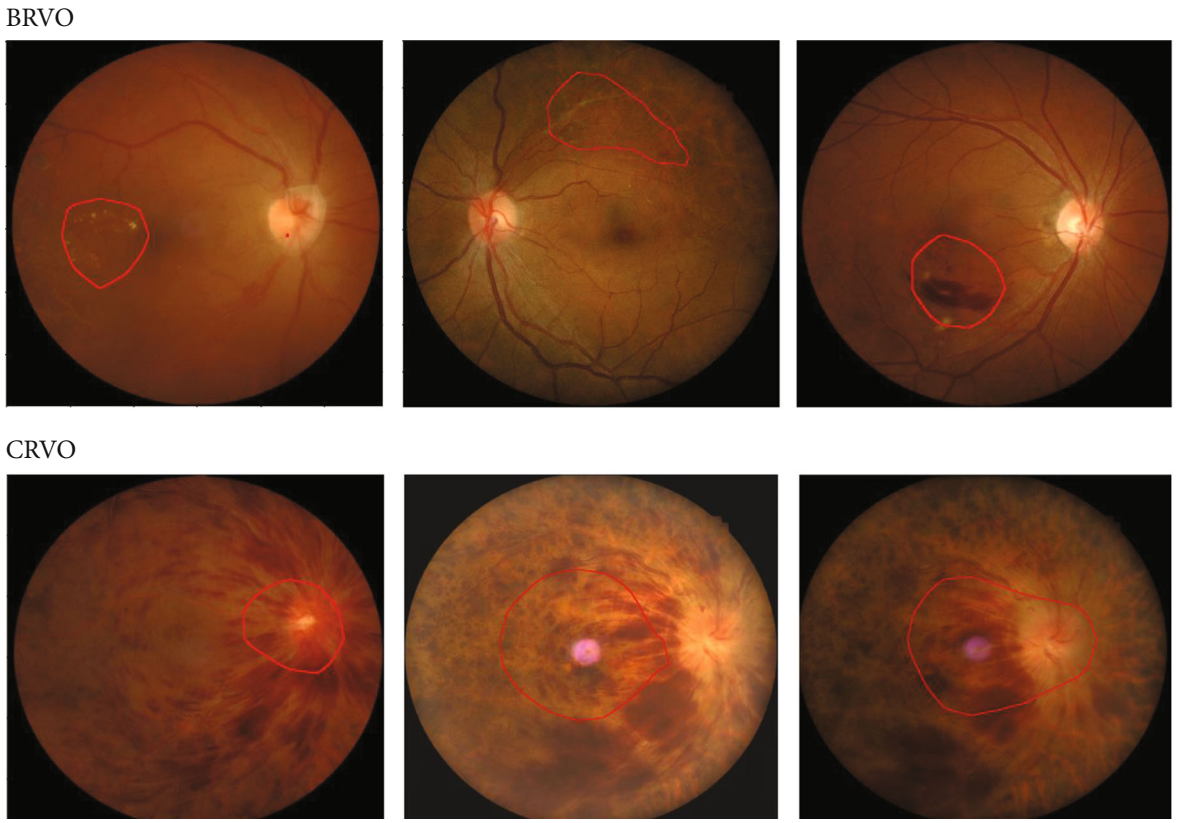


Figure 7: Lesion detection by VGG-CAM network on problematic areas of BRVO and CRVO.

indicates the central area of hemorrhage, namely, lamina cribrosa.

3.3. RVO Classification. The performance of the VGG-CAM network on the validation set is shown in Figure 8 and Table 1. Model scores of the VGG-CAM network (see Table 1) show that the model has a high sensitivity and specificity in classifying BRVO, CRVO, and normal CFIs. When distinguishing between RVO and normal CFIs, the model has only one misclassified image. When distinguishing

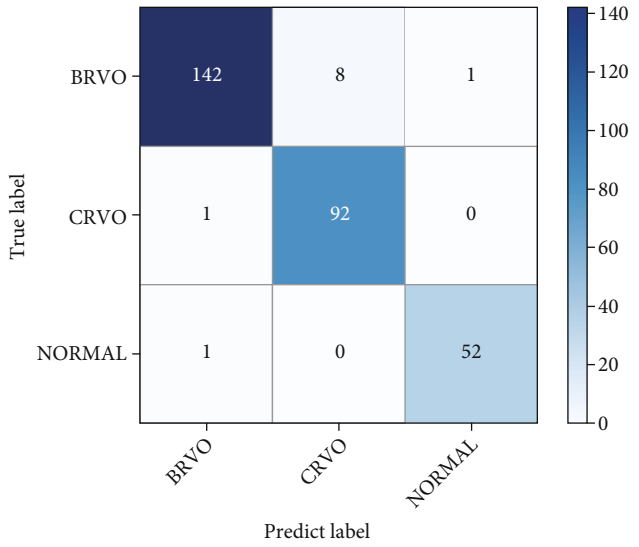


Figure 8: Confusion matrix of VGG-CAM network on validation set.

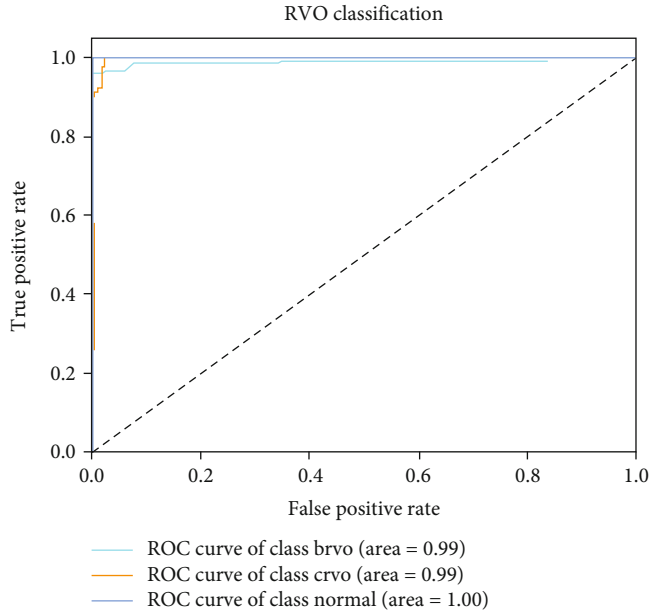


Figure 9: ROC curve of VGG-CAM network performance on RVO classification.

Table 1: VGG-CAM network model scores on RVO classification.

Prediction	Sensitivity	Specificity	Kappa	Number of CFIs
BRVO	0.94	0.99	0.97	151
CRVO	0.99	0.96	0.88	93
Normal	0.98	0.99	0.98	53

Table 2: Comparison of the results of various methods on RVO classification.

Model	Sensitivity	Specificity
Resnet-34	0.92	0.92
Inception-V3	0.90	0.91
MobileNet	0.89	0.90
VGG-CAM without CAM attention	0.95	0.94
VGG-CAM attention	0.97	0.96

between BRVO and CRVO CFIs, it mislabeled eight CRVO images as BRVO. However, the sensitivity and specificity of classifying the three labels are above 94%, and the specificity is above 96%. The results of BRVO and normal CFIs are over 97% in the Kappa coefficient, but the results of CRVO CFIs are only 88%.

From the experimental results in Table 2, it can be seen that the sensitivity and specificity of the model after adding CAM and CAM attention layers are significantly improved when compared with the current classification models with better effects. From the results with CAM attention and without CAM attention, it can be seen that the CAM attention layer enables the model to more effectively extract lesion areas to enhance the final classification effect.

The following ROC curve (see Figure 9) plots the false positive rate (FPR) against the true position rate (TPR). The closer the curve is to (0,1), the more sensitive and accurate the model is. It shows that the area of all curves in the VGG-CAM model reaches 0.99, where the area under the curve of the normal label (1.00) indicates the model is capable of distinguishing between normal CFIs and RVO CFIs. The curves of the BRVO and CRVO labels have an area of 0.99, which indicates a slight probability of mislabeling between each other.

4. Conclusions

This study proposes a hybrid CNN, VGG-CAM, for RVO classification and lesion detection. The CAM attention layer was introduced to enhance the model's attention to the lesion area, and the network parameters learned from the ultralarge data set were used for the initialization of this network by migration learning. Stage training was used to reduce the training time of the model and improve the parameter optimization ability. Further, based on unsupervised learning method, the global average pooling and class activation methods were also used for lesion detection. The experimental results showed that the proposed model can accurately classify BRVO, CRVO, and normal fundus images, detect the lesion areas, and give the prediction results and clinical basis for the resulting judgment.

However, BRVO did not perform as well as CRVO and normal CFIs in sensitivity. For CRVO, the current lesion detection branches cannot achieve a high-precision prediction. This proposed model was only used for the preliminary study of fundus images in the field of view with a 55-degree lens. We concluded that the samples lacked diversity under the specific shooting field of the equipment. In future works, we will improve the model performance with image data from different medical devices and different fields of view

and further improve the lesion detection accuracy of the model through a supervised learning method.

Data Availability

The underlying data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

Guanghua Zhang, Bin Sun, and Zhaoxia Zhang contributed to the work equally and should be regarded as co-first authors.

Acknowledgments

This work was supported by the Research Funds of Shanxi Transformation and Comprehensive Reform Demonstration Zone (Grant No. 2018KJCX04), the Fund for Shanxi "1331 Project", and the Key Research and Development Program of Shanxi Province (Grant No. 201903D311009). The work was also partially sponsored by the Scientific Innovation Plan of Universities in Shanxi Province (Grant No. 2021L575), the Innovation and Entrepreneurship Training Program for College Students (Grant No. 202111242006X), the Shanxi Scholarship Council of China (Grant No. 2020-149), the Shenzhen Fund for Guangdong Provincial High-level Clinical Key Specialties (SZGSP014), the Sanming Project of Medicine in Shenzhen (SZSM202011015), and the Shenzhen Science and Technology Planning Project (KCXFZ20211020163813019).

References

[1] J. Yang, X. Dong, Y. Hu et al., "Fully automatic arteriovenous segmentation in retinal images via topology-aware generative adversarial networks," *Interdisciplinary Sciences, Computational Life Sciences*, vol. 12, no. 3, 2020.

[2] B. Horton William, A. Jahn Linda, M. Hartline Lee, W. Aylor Kevin, T. Patrie James, and E. J. Barrett, "Hyperglycemia does not inhibit Insulin's effects on microvascular perfusion in healthy humans: a randomized crossover study," *American Journal of Physiology Endocrinology and Metabolism*, vol. 319, no. 4, pp. E753–E762, 2020.

[3] W. Shiming and B. Xianyi, "Hyperlipidemia, blood lipid level, and the risk of glaucoma: a meta-analysis," *Investigative Ophthalmology & Visual Science*, vol. 60, no. 4, p. 1028, 2019.

[4] D. Jiajun, Z. Guoning, Y. Yanhong, L. Miao, and H. Yao, "Blood pressure and hypertension prevalence among oldest-old in China for 16 year: based on CLHLS," *BMC Geriatrics*, vol. 19, no. 1, 2019.

[5] M. Perry, "Eye disease in older adults: risk factors and treatments," *Journal of Community Nursing*, vol. 34, no. 3, 2020.

[6] J. Xu, J. Shen, C. Wan, Q. Jiang, Z. Yan, and W. Yang, "A few-shot learning-based retinal vessel segmentation method for assisting in the central serous chorioretinopathy laser surgery," *Frontiers in Medicine*, vol. 9, article 821565, 2022.

[7] J. Zhao, Y. Lu, Y. Qian, Y. Luo, and W. Yang, "Emerging trends and research foci in artificial intelligence for retinal diseases: bibliometric and Visualization study," *Journal of Medical Internet Research*, vol. 24, no. 6, article e37532, 2022.

[8] N. Luke, J. Talks Stephen, A. Winfried, T. Katherine, and S. Sobha, "Retinal vein occlusion (RVO) guideline: executive summary," *Eye*, vol. 36, no. 5, pp. 909–912, 2022.

[9] S. Cugati, J. J. Wang, E. Rochtchina, and P. Mitchell, "Ten-year incidence of retinal vein occlusion in an older population," *Archives of Ophthalmology*, vol. 124, no. 5, pp. 726–732, 2006.

[10] M. P. Sindhu, "Exudate extraction from fundus images using machine learning," *International Journal of Biomedical and Clinical Engineering*, vol. 11, no. 1, pp. 1–16, 2022.

[11] Y. S. Goker, C. U. Atılgan, K. Tekin et al., "Association between disorganization of the retinal inner layers and capillary non-perfusion area in patients with retinal vein occlusion," *Arquivos Brasileiros de Oftalmologia*, vol. 83, pp. 497–504, 2020.

[12] S. Alqadri, M. M. Adil, M. Watanabe, and A. I. Qureshi, "Patterns of collateral formation in basilar artery steno-occlusive diseases," *Journal of Vascular and Interventional Neurology*, vol. 6, no. 2, pp. 9–13, 2013.

[13] R. Murugan and P. Roy, "MicroNet: microaneurysm detection in retinal fundus images using convolutional neural network," *Soft Computing*, vol. 26, no. 3, pp. 1057–1066, 2022.

[14] E. A. Hamed, H. A. Elwakeel, H. A. Eldin, Y. M. BedierElkiran, and M. F. Kamel, "Chemical and mechanochemical catheter-directed sclerotherapy in varicose vein ablation," *The Egyptian Journal of Surgery*, vol. 40, no. 1, pp. 90–98, 2021.

[15] M. Suzani and A. T. Moore, "Coats disease: fluorescein angiography guided management," *Acta Ophthalmologica*, vol. 90, 2012.

[16] T. Wu, L. Liting, Z. Tianer, and W. Xuesen, "Deep learning-based risk classification and auxiliary diagnosis of macular edema," *Intelligence-Based Medicine*, vol. 6, article 100053, 2022.

[17] S. R. Singh, D. C. Parameswarappa, V. Govindahari, M. Lupidi, and J. Chhablani, "Clinical and angiographic characterization of choroidal neovascularization in diabetic retinopathy," *European Journal of Ophthalmology*, vol. 31, no. 2, pp. 584–591, 2021.

[18] J. Shilpa and P. T. Karule, "Haemorrhages detection using geometrical techniques," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 8, no. 4, pp. 436–445, 2020.

[19] G. T. Frangieh, W. R. Green, E. Barraquer-Somers, and D. Finkelstein, "Histopathologic study of nine branch retinal vein occlusions," *Archives of Ophthalmology*, vol. 100, no. 7, pp. 1132–1140, 1982.

[20] W. R. Green, C. C. Chan, G. M. Hutchins, and J. M. Terry, "Central retinal vein occlusion: a prospective histopathologic study of 29 eyes in 28 cases," *Retina*, vol. 1, no. 1, pp. 27–55, 1981.

[21] P. Song, Y. Xu, M. Zha, Y. Zhang, and I. Rudan, "Global epidemiology of retinal vein occlusion: a systematic review and meta-analysis of prevalence, incidence, and risk factors," *Journal of Global Health*, vol. 9, no. 1, article 010427, 2019.

[22] M. Laouri, E. Chen, M. Looman, and M. Gallagher, "The burden of disease of retinal vein occlusion: review of the literature," *Eye*, vol. 25, no. 8, pp. 981–988, 2011.

[23] S. Sivaprasad, W. M. Amoaku, P. Hykin, and RVO Guideline Group, "The Royal College of Ophthalmologists guidelines on retinal vein occlusions: executive summary," *Eye*, vol. 29, no. 12, pp. 1633–1638, 2015.

[24] M. Sun, X. F. Hao, L. K. Xie, Q. Jin, S. H. Wang, and J. Xu, "Methods for making animal models of retinal vein occlusion," *International Eye Research*, vol. 2, no. 3, pp. 164–169, 2021.

[25] I. L. McAllister, "Central retinal vein occlusion: a review," *Clinical & Experimental Ophthalmology*, vol. 40, no. 1, pp. 48–58, 2012.

[26] S. Suthaharan, "Machine Learning Models and Algorithms for Big Data Classification," in *Integrated Series in Information Systems*, pp. 289–307, Springer, New York, NY USA, 2016.

[27] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[28] Y. LeCun, B. Boser, J. S. Denker et al., "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, pp. 84–90, 2017.

[30] M. Voets, K. Møllersen, and L. A. Bongo, "Replication study: development and validation of deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," 2018, arXiv preprint arXiv:1803.04337.

[31] H. Zhang, Z. Chen, Z. Chi, and H. Fu, "Hierarchical local binary pattern for branch retinal vein occlusion recognition with fluorescein angiography images," *Electronics Letters*, vol. 50, no. 25, pp. 1902–1904, 2014.

[32] R. Zhao, Z. Chen, and Z. Chi, "Convolutional neural networks for branch retinal vein occlusion recognition," in *Proceeding of the 2015 IEEE International Conference on Information and Automation*, pp. 1633–1636, Lijiang, China, 2015.

[33] E. Zahra, B. Ali, and W. Siddique, "Medical image segmentation using a U-Net type of architecture," 2020, arXiv preprint arXiv:2005.05218.

[34] C. Wan, X. Zhou, Q. You et al., "Retinal image enhancement using cycle-constraint adversarial network," *Frontiers in Medicine*, vol. 8, article 793726, 2022.

[35] Q. Chen, W. H. Yu, S. Lin et al., "Artificial intelligence can assist with diagnosing retinal vein occlusion," *International Journal of Ophthalmology*, vol. 14, no. 12, pp. 1895–1902, 2021.

[36] S. Heung-II, *Deep Learning for Medical Analysis*, K. Zhou, Ed., Elsevier Inc, 2017.

[37] P. Li, N. Yi, C. Ding, L. I. Sheng, and H. Min, "Research on classification diagnosis model of psoriasis based on deep residual," *Digital Chinese Medicine*, vol. 4, no. 2, pp. 92–101, 2021.

[38] C. Jianfang, Y. Minmin, J. Yiming, T. Xiaodong, and Z. Zibang, "Application of a modified Inception-v3 model in the dynasty-based classification of ancient murals," *EURASIP Journal on Advances in Signal Processing*, vol. 2021, no. 1, 2021.

[39] A. Michele, V. Colin, and D. D. Santika, "MobileNet Convolutional Neural Networks and Support Vector Machines for Palmprint Recognition," *Procedia Computer Science*, vol. 157, pp. 110–117, 2019.

[40] M. Sharma, B. Jain, C. Kargeti, V. Gupta, and D. Gupta, "Detection and diagnosis of skin diseases using residual neural networks (RESNET)," *International Journal of Image and Graphics*, vol. 21, no. 5, 2021.

[41] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, Las Vegas, NV, USA, 2016.

[42] A. G. Howard, M. Zhu, B. Chen et al., "MobileNets: efficient convolutional neural networks for mobile vision applications," *Computer Vision and Pattern Recognition*, vol. 1, pp. 1–9, 2017.

[43] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, arXiv preprint arXiv: 1409.1556.

[44] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, arXiv preprint arXiv:1312.4400.

[45] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *Journal of Vlsi Signal Processing Systems for Signal, Image and Video Technology*, vol. 38, no. 1, pp. 35–44, 2004.

[46] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," *Advances in Neural Information Processing Systems*, vol. 27, pp. 3320–3328, 2014.

[47] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[48] C.-Y. Tai, W. Meng-Ru, Y.-W. Chu, and S.-Y. Chu, "GraphSW: a training protocol based on stage-wise training for GNN-based recommender model," 2019, arXiv preprint arXiv:1908.05611.

[49] E. Barshan and P. Fieguth, "Stage-wise training: an improved feature learning strategy for deep models," *Feature Extraction: Modern Questions and Challenges*, vol. 44, pp. 49–59, 2015.