

Automated Detection of Central Retinal Vein Occlusion Using Convolutional Neural Network

Bismita Choudhury, Patrick H.H. Then, and Valliappan Raman

Abstract The Central Retinal Vein Occlusion (CRVO) is the next supreme reason for the vision loss among the elderly people, after the Diabetic Retinopathy. The CRVO causes abrupt, painless vision loss in the eye that can lead to visual impairment over the time. Therefore, the early diagnosis of CRVO is very important to prevent the complications related to CRVO. But, the early symptoms of CRVO are so subtle that manually observing those signs in the retina image by the ophthalmologists is difficult and time consuming process. There are automatic detection systems for diagnosing ocular disease, but their performance depends on various factors. The haemorrhages, the early sign of CRVO, can be of different size, color and texture from dot haemorrhages to flame shaped. For reliable detection of the haemorrhages of all types; multifaceted pattern recognition techniques are required. To analyse the tortuosity and dilation of the veins, complex mathematical analysis is required in order to extract those features. Moreover, the performance of such feature extraction methods and automatic detection system depends on the quality of the acquired image. In this chapter, we have proposed a prototype for automated detection of the CRVO using the deep learning approach. We have designed a Convolutional Neural Network (CNN) to recognize the retina with CRVO. The advantage of using CNN is that no extra feature extraction step is required. We have trained the CNN to learn the features from the retina images having CRVO and classify them from the normal retina image. We have obtained an accuracy of 97.56% for the recognition of CRVO.

Keywords Retinal vein occlusion • Central retinal vein occlusion • Convolution • Features

B. Choudhury (✉) • P.H.H. Then • V. Raman
Centre for Digital Futures and Faculty of Engineering, Computing and Science, Swinburne
University of Technology, Sarawak Campus, Kuching, Sarawak, Malaysia
e-mail: bismi.choudhury@gmail.com

© Springer International Publishing AG 2017
S.C. Suh, T. Anthony (eds.), *Big Data and Visual Analytics*,
https://doi.org/10.1007/978-3-319-63917-8_1

1 Introduction

The retina is the light sensitive tissue covering the interior surface of the eye. The cornea and the lens focus light rays on the retina. Then, the retina transforms the light received into the electrical impulses and sends to the brain via the optic nerve. Thereby, a person interprets those impulses as images. The cornea and the lens in the eye behave like the camera lens, while the retina is analogous to the film. Figure 1 shows the retina image and its different features like optic disc, fovea, macula and blood vessels.

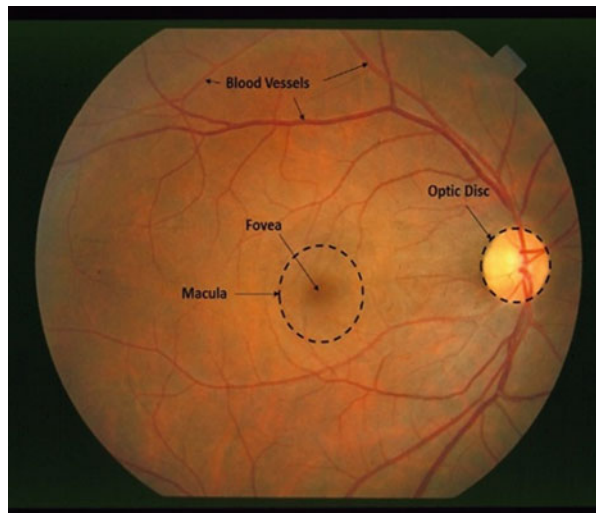
The Retinal Vein Occlusion (RVO) is an obstruction of the small blood carrying veins those drain out the blood from the retina. There are one major artery, called the central retinal artery, and one major vein, called the central retinal vein, in the retina. The Central Retinal Vein Occlusion (CRVO) occurs when a thrombosis is formed in this vein and causes leaking of blood and excess fluid into the retina. This fluid often accumulates around the macula, the region for the central vision, in the retina. Sometimes the blockage occurs when the veins in the eye are too narrow [1].

The diagnostic criteria for CRVO are characterized by flame-shaped, dot or punctate retinal hemorrhages or both in all four quadrants of the retina, dilated and tortuous retinal veins, and optic disc swelling [2].

The CRVO can be either ischemic or non-ischemic. About 75% of the cases, non-ischemic CRVO is a less severe form of CRVO and usually has a chance for better visual acuity. Ischemic CRVO is a very severe stage of CRVO where significant complications arise and can lead to the vision loss and probably damage the eye [1].

Macular edema is the prime reason for the vision loss in CRVO. The fluid accumulated in the macular area of the retina causes swelling or edema of the macula. It causes the central vision of a person to become blurry. The patients

Fig. 1 Retina and its different features



with the macular edema following CRVO might have some of the most common symptoms, such as blurred vision, distorted vision, or vision loss in all or part of the eye [2].

The lack of oxygen (ischemia) in the retina can lead to the growth of the abnormal blood vessels. The patients with ischemic CRVO develop neovascular glaucoma over three months or longer period of time. In neovascular glaucoma, the abnormal blood vessels increase the pressure in the eye that can cause severe pain and vision loss [3].

Usually, the people who are aged 50 and older have higher chance of suffering from CRVO. The probability of occurring CRVO is higher in people having diabetes, high blood pressure, high cholesterol, or other health issues that interrupt blood flow. The symptoms of retinal vein occlusion can be range from indistinct to very distinct. Most of the time, just one eye suffers from painless blurring or loss of vision. Initially, the blurring or the vision loss of the eye might be minor, but this situation gets worse over the next few hours or days. Sometimes the patients might lose the complete vision almost immediately. In 6–17% of the cases, the second eye also develops the vein occlusion after the first one. Up to 34% of eyes with non-ischemic CRVO convert to ischemic CRVO over 3 years [1].

It is crucial to recognize CRVO to prevent further damage in the eye due to vein occlusion and treat all the possible risk factors to minimize the risk of the other eye to form CRVO. The risk factor of CRVO includes hypertension, diabetes, hyperlipidemia, blood hyperviscosity, vascular cerebral stroke and thrombophilia. The treatment of any of these risk factors reduces the risk of a further vein occlusion occurring in either eye. It may also help to reduce the risk of another blood vessel blockage, such as may happen in a stroke (affecting the brain) or a heart attack or, in those with rare blood disorders, a blocked vein in the leg (deep vein thrombosis) or lung (pulmonary embolism). There is no cure, but early treatment may improve vision or keep the vision from worsening [4].

The automatic detection of CRVO in the early stage can prevent the total vision loss. The automatic detection can save lots of time for the ophthalmologist. Rather than putting lots of effort in diagnosis, they can put more time and effort for the treatment. Thereby, the patients will receive the treatment as early as possible. It will be also beneficial for the hospitals and the patients in terms of saving time and money. For the diagnosis of retinal disease, mostly fluorescein angiographic image or color fundus image is taken. However, compared to angiographic images color fundus images are more popular in the literature of automatic retina analysis. Color fundus images are widely used because it is inexpensive, non-invasive, can store for future reference and ophthalmologists can examine those images in real time irrespective of time and location. In all the computer aided detection system, the abnormal lesions or features are detected to recognize the particular disease.

The early symptoms of CRVO are very subtle to detect. When non-ischemic CRVO forms, the retina remains moderately normal. So, there is higher chance that general automatic detection system fails to detect CRVO in the earliest stage.

Another clinical sign for CRVO is dilated tortuous vein. So, it is important to segment and analyse the retinal vasculature. Most importantly, it is required to detect the vein and calculate the tortuosity index and analyse the change in blood vessels due to dilation. Moreover, it is crucial to detect any newly generated blood vessels leading to neovascularization for ischemic CRVO. Another clinical characteristic of CRVO is haemorrhages, which can be of different size, color and texture. The haemorrhages in CRVO are mostly dot haemorrhage and flame shaped haemorrhage. The dot haemorrhages appear similar to the microaneurysms. Therefore, the segmentation process can't distinguish between the dot haemorrhage and the microaneurysm. Therefore, by default the dot haemorrhages are detected as microaneurysm by the automated microaneurysm detection process. The literature does not support much description about the automated detection of the haemorrhages [5]. In the ischemic stage, there will be multiple Cotton wool spots and the literature doesn't provide much attention to the automatic detection of cotton wool spots. In short, the problem with the automatic detection of the CRVO is that, the sophisticated segmentation and feature extraction algorithms are required for each of the clinical signs. For example, for reliable detection of the haemorrhages of all types; multifaceted pattern recognition techniques are required. For analysing dilated veins, tortuous vein, newly formed blood vessels, we need complicated mathematical approach for such feature extraction. Again, the performance of such algorithms and the classification depends on the image quality of the retina image acquired. The inter-and intra-image contrast, luminosity and color variability present in the images make it challenging to detect these abnormal features. To our best knowledge, no research work related to automatic detection of CRVO has been done yet.

In this chapter, we approached the deep learning method for detecting the CRVO. We have exploited the architecture of Convolutional Neural Network (CNN) and designed a new network to recognize CRVO. The advantage of using CNN is that, design of complex, sophisticated feature extraction algorithms for all the clinical signs of CRVO are not necessary. The convolution layer in the neural network extracts the features by itself. Moreover, CNN takes care of image size, quality etc. The chapter is organized as follows: the first part will briefly describe about the types of CRVO. In the second section, the computer aided detection system for medical images will be discussed. In the third section, we will review previous related work on the automated detection of vein occlusion. The fourth section will describe the theory and architecture of the Convolutional Neural Network. The fifth section will describe the design of the CNN for the recognition of CRVO.

2 Central Retinal Vein Occlusion (CRVO)

The two types of CRVO, ischemic and non-ischemic, have very different diagnoses and management criteria from each other. Both the types are briefly discussed below:

2.1 *Non-Ischemic CRVO*

It was reported that majority of the cases (about 70%), the patients suffer from non-ischemic CRVO [3]. Over 3 years, 34% of non-ischemic CRVO eyes progressed to ischemic CRVO. There is low risk of neovascularization in case of non-ischemic CRVO. The clinical features of non-ischemic CRVO are as follows:

- Vision acuity $>20/200$.
- The low risk of forming neovascularization.
- More dot & blot hemorrhages.
- The retina in non-ischemic CRVO will be moderately normal.
- There is no Afferent Pupillary Defect (APD).

Figure 2 shows the retina image with non-ischemic CRVO.

2.2 *Ischemic CRVO*

According to the fluorescein angiographic evidence, the ischemic CRVO is defined as of more than 10 disc areas of capillary non-perfusion on seven-field fundus fluorescein angiography [1]. It is associated with an increased risk of neovascularization and has a worse prognosis [3, 6]. There is a 30% chance of converting non-ischemic to ischemic CRVO [1]. More than 90% of patients with ischemic CRVO have a final visual acuity of 6/60 or worse [6]. The clinical features of ischemic CRVO are as follows:

- Visual acuity $<20/200$
- The high risk of forming neovascularization
- Widespread superficial hemorrhages.

Fig. 2 Non-ischemic CRVO



Fig. 3 Ischemic CRVO

- Multiple Cotton Wool Spots.
- Poor capillary perfusion (ten or more cotton wool spots or ten DD capillary non-perfusion on fluorescein angiography).
- Turbid, orange, edematous retina.
- Poor prognosis
- Degree of retinal vein dilatation and tortuosity.
- High Relative Afferent Pupillary Defect (+RAPD).

Figure 3 shows the retina image with ischemic CRVO.

3 Computer Aided Detection (CAD)

The Computer Aided Detection (CAD) systems are designed to assist physicians in the evaluation of medical images. CAD is rapidly growing in the field of radiology to improve the accuracy and consistency of the radiologists' image interpretation. CAD system processes digital images and highlight the suspicious section to evaluate the possible disease. The goal of the CAD systems is to detect the earliest signs of abnormality in the patients' medical image that human professionals cannot. It is pattern recognition software that automatically detects the suspicious features in the image to get the attention from the radiologist and reduce the false negative reading. The computer algorithm for automatic detection usually consists of multiple steps, including image processing, image feature extraction, and data classification through a different classifier such as artificial neural networks (ANN) [7]. The CAD systems are being used for different image modalities from Magnetic Resonance Imaging (MRI), Computed Tomography (CT), ultrasound imaging, retinal funduscopy image etc.

Computerized scheme for CAD has three main components involving three different technologies [7]. Those are:

1. *Image processing and segmentation:* In this component, the medical image is enhanced to roughly detect/extract the candidates for suspicious lesions or patterns. There are various image enhancement techniques for different lesions. Some of the commonly used techniques are Fourier analysis, morphological filters, wave analysis, different image analysis techniques and artificial neural network (ANN).
2. *Feature extraction:* The different image features are quantized in terms of size, shape and contrast from the selected candidates from the first step. It is possible to define multiple features using mathematical formulas. Initially, the CAD might be fed into the physicians' knowledge based on their observations. One of the important factors in developing CAD is to distinguish abnormal features or lesions from the normal structures.
3. *Classification:* The data is analyzed to differentiate the normal and abnormal patterns based on the extracted features. A rule based method can be applied from the understanding of normal and abnormal lesions. Other than the rule based approach, discriminant analysis, neural network and decision tree can be used.

Figure 4 shows the block diagram of the general CAD system.

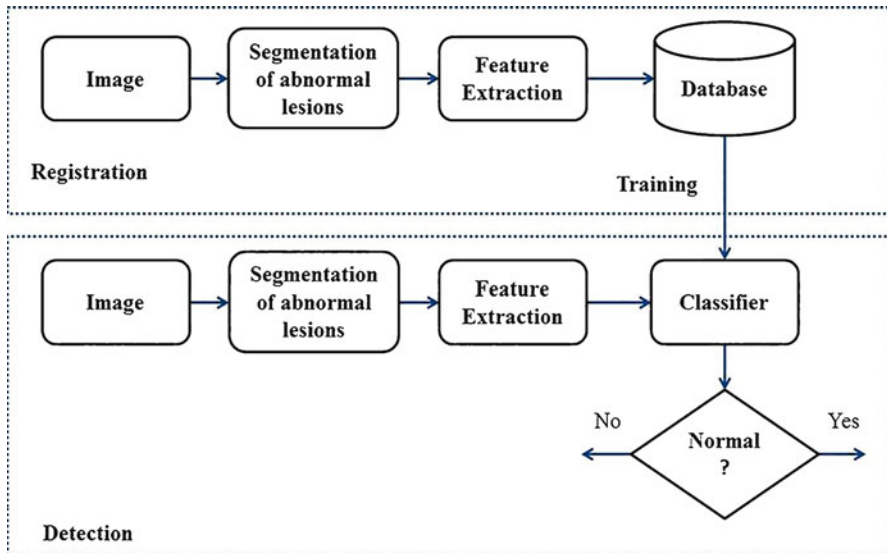


Fig. 4 Block diagram of the general computer aided detection system

4 State of the Art for RVO Detection

There are a few research works available for automatic detection of retinal vein occlusion. The existing approaches for automated detection of RVO are discussed below:

4.1 Feature Representation Techniques

Zhang et al. in [8] proposed a Hierarchical Local Binary Pattern (HLBP) to represent the features of Branch Retinal Vein Occlusion (BRVO), which is another type of vein occlusion in the distal branches [14]. They provided a hierarchical combination of Linear Binary Pattern (LBP)-coding and max pooling inspired by the convolutional neural network. There are two levels in HLBP and each level consists of a max-pooling layer and an LBP-coding layer. For the HLBP calculation, first, max-pooling is performed on the Fluorescein Angiography (FA) image to generate a feature map in the first level. Then, the LBP is performed on the feature map generated from level 1 and generates an LBP1 feature map. Secondly, in the second level, max pooling is performed on the LBP1 feature map producing another feature map and an LBP2 map respectively. Finally, a feature vector of the FA image is generated by combining the histograms of the LBP1 map and the LBP2. They used SVM classifier with the linear kernel for the classification and achieved a mean accuracy of 96.1%.

Gayathri et al. in [9] presented a feature representation scheme to diagnosis RVO. The textures of the blood vessel are extracted using Completed LBP technique. CLBP is represented by its center pixel and a local difference sign-magnitude transforms (LDSMT). To find the two peaks in the histogram, global thresholding is used. The center pixels are coded by binary codes which are termed as CLBP_CENTER (CLBP_C). Then, the image is divided by LDMST into two components: CLBP-Sign and CLBP-Magnitude. They are, then, combined to produce the histogram of the image. The texture image is first constructed from the histogram of the image. Then, the neural network is trained to classify the extracted features and identify the affected retinal images.

4.2 Fractal Analysis

Fazekas et al. in [10] compared two blood vessel segmentation methods. Then fractal properties of blood vessels are analyzed to distinguish normal retina image and RVO images. Both blood vessel segmentation methods are based on directional response vector similarity and the region growing. The first method yields a binary map of the retinal blood vessels of the input retinal fundus image. This method used hysteresis thresholding to apply the region growing procedure to the response vector

similarities of neighboring pixels within the fundus image. In the second method, Gabor functions are used as template matching procedure to calculate the response. The fractal analysis was performed on a number of retinal images by combining the segmented images and their skeletonized versions. In fractal analysis, the box-dimension is used to estimate the fractal dimension via box counting. The lower and upper box counting dimensions of a subset, respectively, are defined as follows:

$$\underline{dim}_B(F) = \lim_{r \rightarrow 0} \frac{\log N_r(F)}{-\log r}; \overline{dim}_B(F) = \overline{\lim}_{r \rightarrow 0} \frac{\log N_r(F)}{-\log r} \quad (1)$$

If the lower and upper box-counting dimensions are equal, then their common value is referred to as the box-counting dimension of F and is denoted with

$$dim_B(F) = \lim_{r \rightarrow 0} \frac{\log N_r(F)}{-\log r} \quad (2)$$

where $N_r(F)$ can be of the following:” (1) the smallest number of closed balls (i.e., disks, spheres) of radius $r > 0$ that cover F; (2) the smallest number of cubes of side r that cover F; (3) the number of r -mesh cubes that intersect F; (4) the smallest number of sets of diameter at most r that cover F; (5) the largest number of disjoint balls of radius r with centers in F.”

The fractal dimension is calculated for both the skeletonized images of normal retina and the retina with RVO. There is no significant difference in the fractal dimension of healthy eyes. But, the fractal dimension is quite visible in case of retina with CRVO. The fractal dimensions computed seemed to be beneficial in separating the different types of RVO.

4.3 Deep Learning Approach

Zhao et al. in [11] proposed a patch based and an image based voting method for the recognition of BRVO. They exploited Convolutional Neural Network (CNN) to classify the normal and BRVO color fundus images. They extracted the green channel of the color fundus image and performed image preprocessing to improve the image quality. In the patch based method they divided the whole image into small patches and put labels on each patches to train the CNN. If the patch has BRVO features, labeled as BRVO otherwise labeled as normal. During the training phase, only the patches with the obvious BRVO feature are labeled as BRVO. Those ambiguous patches are discarded. The testing is done by feeding all the patches of a test image to the trained CNN. They kept the threshold of 15 patches for each test. If the test image passes the threshold, the testing image is classified to BRVO. In the image based scheme, at first, three operations, noise adding, flipping, and rotation are performed on a preprocessed image. Depending on the classification results of these four images, the final decision for a test image is made. “If the classification

Table 1 Summary of the research works done for identifying RVO

Author	Year	Type of RVO	Method	Accuracy
Zhang et al. [8]	2015	BRVO	Hierarchical linear binary pattern, support vector machine	96.1%
Gayathri et al. [9]	2014	–	Complete linear binary pattern, neural network	–
Fazekas et al. [10]	2015	CRVO, BRVO, HRVO	Fractal properties of blood vessels	–
Zhao et al. [11]	2014	BRVO	Convolutional neural network	97%

results of the three new images (noisy, flipped, and rotated) are the same, the test image is classified in the class of the three new images. Otherwise, the test image is classified to the class of the original image”. For patch based method they obtained 98.5% and for image based method they obtained 97%. Compared to patch based method, image based voting method is more practical.

The Table 1 summarize the various research done for the automated diagnosis of retinal vein occlusion

From the limited previous research work, it is clear that less attention has been paid towards the automatic detection of RVO. The fractal analysis described in [10] provided the calculation of fractal dimension of RVO images and the possibility of using those values for quantifying the different types of RVO. No clear information is provided regarding the accuracy of the methods described in [9, 10]. In [9, 10], retinal vascular structure of the color fundus image is analysed to extract the features and used classifier to detect RVO. In [8], the features are extracted from the whole image to detect BRVO in Fluorescein Angiography (FA) image. In [11], CNN is used to detect the BRVO in color fundus image. The majority of the available research works are on automatic detection of BRVO. No research work has been found that focuses on detecting haemorrhages, analyse vessel tortuosity and dilation to recognize RVO. There is no existing method for the automatic detection of CRVO considering the fact that the visual imparity is more severe in case of CRVO. Hence, it is very important to design an automatic detection system for recognizing CRVO. For automatic detection of CRVO there can be two approaches. One, individually extract the abnormal features from the segmented retinal pathologies by compound pattern recognition techniques and then, fed them to a classifier to identify the CRVO. Otherwise, extract the abnormal features from the whole image and use supervised machine learning classifier to identify CRVO.

5 The Proposed Methodology

We opted a deep learning approach to extract the features from the raw retina image and classify as CRVO. We explored the architecture of the Convolutional Neural Network (CNN) and designed a CNN for learning the CRVO

features and classify the CRVO images from the normal retina images. There are some advantages of using CNN. First, we do not have to design individual segmentation and feature extraction algorithms for all the clinical signs of CRVO (haemorrhages, cotton wool spots, dilated tortuous veins and newly formed blood vessels) and rely on the accuracy of such algorithms to classify CRVO. Second, the convolutional neural network is invariant to any kind of distortion in the image, for example, the different lighting condition, camera position, partial occlusion etc. Third, easier to train compared to conventional Neural Network (NN) due to reduced parameters used during training. Fourth, memory requirement is less as convolution layer use same parameters to extract the features across the different locations in an image. In this study, we collected the normal retina image and retina with CRVO image from multiple publicly available databases. We used STARE database (<http://cecas.clemson.edu/~ahoover/stare/>), DRIVE database (<http://www.isi.uu.nl/Research/Databases/DRIVE/>), dataset of Dr. Hossein Rabbani (<https://sites.google.com/site/hosseinrabbanikhorasgani/datasets-1>) and Retina Image Bank (<http://imagebank.asrs.org/discover-new/files/1/25?q>). The existing methods conducted their experiments on different datasets and since those datasets have different image size and quality we cannot compare their performance directly. Because, the experimental results on one database are not consistent for all other different databases. A method showing high performance in one database might not show same high performance in other database. Since, we conducted our experiments on the retina images from various sources and all images are of different size and quality; we can say that the proposed method for CRVO detection is a versatile method whose performance should be consistent for any database of retinal fundus image. Therefore, it is feasible to implement in real.

5.1 The Basic of the Convolutional Neural Network (CNN)

The Convolutional Neural Network is the advanced version of the general Neural Network (NN); used in various areas, including image and pattern recognition, speech recognition, natural language processing, and video analysis [12]. The CNN facilitates the deep learning to extract abstract features from the raw image pixels.

CNNs take a biological inspiration from the visual cortex. The visual cortex has lots of small cells that are sensitive to specific regions of the visual field, called the receptive field. This small group of cells functions as local filters over the input space. This idea was expanded upon by Hubel and Wiesel, where they showed that some individual neuronal cells in the brain responded (or fired) only in the presence of edges of a certain orientation. For example, some neurons fired when exposed to vertical edges and some when shown horizontal or diagonal edges. They found out that all of these neurons were structured as a columnar architecture and are able to produce visual perception [13]. This idea of specialized components inside of a system having specific tasks (the neuronal cells in the visual cortex looking for specific characteristics) is one that machines use as well, and is the basis behind

CNNs. By assembling several different layers in a CNN, complex architectures are constructed for classification problems. The CNN architecture consists of four types of layers: convolution layers, pooling/subsampling layers, non-linear layers, and fully connected layers [13, 15].

5.1.1 The Convolutional Layer

The first layer in a CNN is always a Convolutional Layer. The convolution functions as feature extractor that extracts different features of the input. The first convolution layer extracts the low-level features like edges, lines, and corners. Higher-level layers extract the higher-level features. Suppose, the input is of size $M \times M \times D$ and is convolved with K kernels/filters, each of size $n \times n \times D$ separately. Convolution of an input with one kernel produces one output feature. Therefore, the individual convolution with K kernels produces K features. Starting from top-left corner of the input, each kernel is moved from left to right and top to bottom until the kernel reaches the bottom-right corner. For each stride, element-by-element multiplication is done between $n \times n \times D$ elements of the input and $n \times n \times D$ elements of the kernel on each position of the kernel. So, $n \times n \times D$ multiply-accumulate operations are required to create one element of one output feature [13, 15].

5.1.2 The Pooling Layer

The pooling layer reduces the spatial size of the features. It makes the features robust against noise and distortion. It also reduces the number of parameters and computation. There are two ways for down sampling: max pooling and average pooling. Both the pooling functions divide the input into non-overlapping two dimensional space [12].

5.1.3 Non-Linear layer

The non-linear layer adds non linearity to the network as the real world data are non-linear in nature (<https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>). The rectified linear unit (ReLU) is a nonlinear layer that triggers a certain function to signal distinct identification of likely features on each hidden layer. A ReLU performs the function $y = \max(x, 0)$ keeping the output size same as the input. It also helps to train faster.

5.1.4 Fully Connected Layers

In a CNN, the last final layer is a fully connected layer. It is a Multi-Layer Perceptron (MLP) that uses an activation function to calculate the weighted sum of all the features of the previous layer to classify the data into target classes.

5.2 Methodology

Some of the popular Convolutional Networks are LeNet, AlexNet, ZF Net, GoogLeNet, VGGNet and ResNet. The LeNet is the first successful Convolutional Network used for recognizing digits, zip codes etc. After LeNet, AlexNet came as a deeper version of LeNet which was successfully used for object recognition in large scale. ZF Net is modified version of the AlexNet where the hyper parameters are modified. The GoogLeNet introduced inception module to drastically reduce the number of the parameters. VGGNet is a large deep Convolutional Network with 16 Convolutional and Fully Connected layers. ResNet skipped the fully connected layers and made heavy use of batch normalization. Moreover, some CNNs are fine tuned or the architecture is tweaked for different applications. For e.g. in [16], the authors designed a CNN for facial landmark detection. Again in [17] and [18], the basic CNN is fine tuned to identify different EEG signals. Our designed Convolutional Network is based on LeNet architecture. The general structure of LeNet is as follows:

Input=>Conv=>Pool=>Conv=>Pool=>FC=>ReLu=>FC=>Output

Our designed CNN structure is as follows:

Input=>Conv=>ReLU=>Pool=>Conv=>ReLU=>Pool=>Conv=>Re

LU=>Pool=>FC=>ReLU=>FC=>Output

Before feeding the retina image to the CNN, we performed preprocessing to enhance the quality of the image. Since, we have collected the color fundus image of normal retina and the retina with CRVO images from multiple databases, all the images are of different sizes and of different formats. The images from the STARE databases are of size 700×605 and TIF format. The images from the DRIVE database are 565×585 TIFF images. The images from the Dr. Hossein Rabbani are 1612×1536 JPEG images. The each of the images from the Retina Image Bank is of different sizes and format. We converted all the images to TIF format and resized into a standard image size 60×60 .

5.2.1 Image Preprocessing

After converting all the images into 60×60 TIF format, we extracted the green channel as it provides the distinct visual features of the retina compared to other two (Red and Blue) channels. Then, an Average filter of size 5×5 is applied to remove the noise. After that the contrast of that grayscale retina image is enhanced by applying Contrast-limited adaptive histogram equalization (CLAHE). CLAHE operates on small regions in the image and enhance the contrast of each small region individually. A bilinear interpolation is used to combine the neighboring small regions in order to eliminate artificial boundaries. The contrast of the homogeneous areas can be limited to avoid unwanted noise present in the image. Figure 5a shows the normal image and Fig. 5b shows the green channel of the RGB image, Fig. 5c

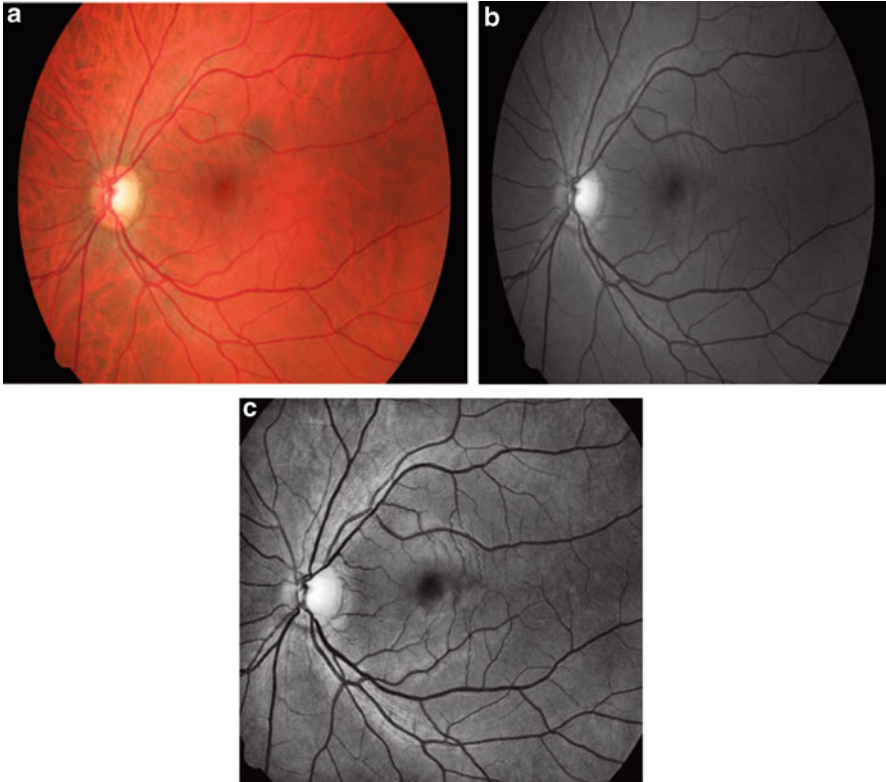


Fig. 5 (a) Normal RGB images, (b) Green channel. (c) Pre-processed image

shows the enhanced image. Figure 6a shows the CRVO image, Fig. 6b shows the green channel and Fig. 6c shows the enhanced image.

5.2.2 The Network Topology

The designed CNN for recognition of CRVO consists of 12 layers, including three convolution layers, three pooling layers, four ReLUs and two fully connected layers. We have two classes: Normal image and CRVO image. The layers in the CNN network are stacked with three sets of convolution layer, followed by ReLU followed by a pooling, followed by a fully connected layer, ReLU and fully connected layer. Finally, the features obtained by the 4th ReLU layer are transferred to the last fully connected layer. Ultimate classification of CRVO is based on these high level features. Softmax function is used as the activation function. The network topology is shown in Fig. 7.

Layer 1: The first convolutional layer convolves the input retina image of size 60×60 . In the first layer, we used 32 filters of size 5×5 with a stride of 1 to outputs

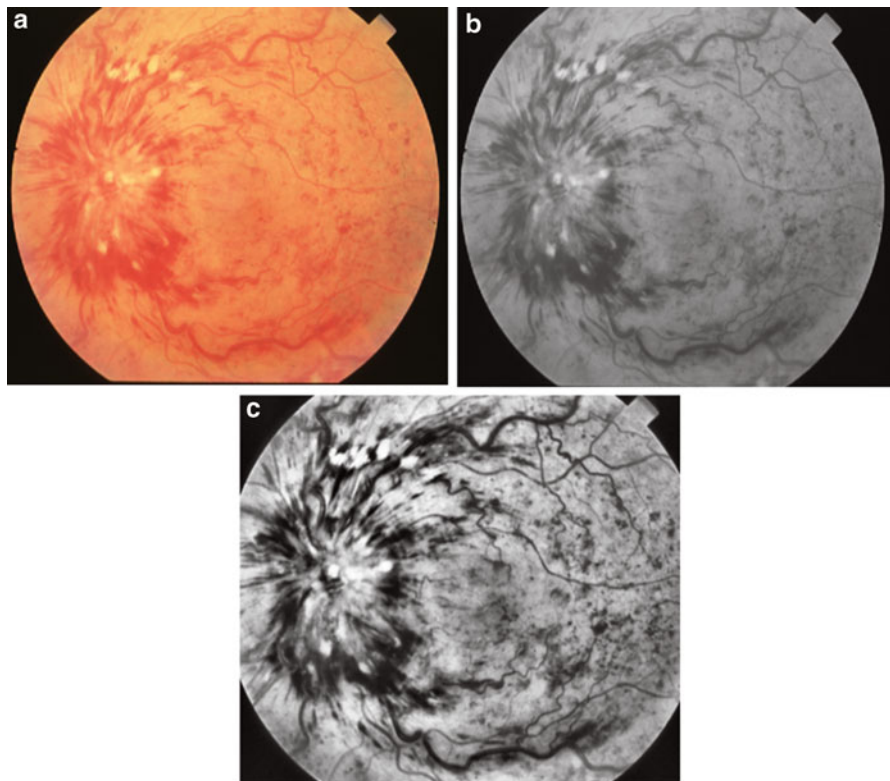


Fig. 6 (a) CRVO image, (b) *Green* channel, (c) Pre-processed image

the feature data map. Mathematically, the operation of a convolution layer can be formulated as follows:

$$y_j^n = f \left(\sum_{i \in N_j} W_{ij}^n y_i^{n-1} - b_j^n \right) \quad (3)$$

where y_i^{n-1} is the input of convolution layer. W_{ij}^n is a convolution kernel weight of layer n with the size of $i \times j$. b_j^n is a bias, y_j^n is the output and N is the number of inputs used to generate y_j^n . After the first convolution layer the output feature map is of size $56 \times 56 \times 32$.

Layer 2: After the first convolution layer, a rectified non-linear unit (ReLU) is used. It increases the nonlinear property keeping the output volume same as the input volume.

Layer 3: The output feature map of ReLU is given as input to a pooling layer. For the pooling layer, we used max pooling to reduce the size of the output feature map and capture the spatial information. A filter size 2×2 and stride 2 are used for the max pooling. The equation of pooling layer can be given by,

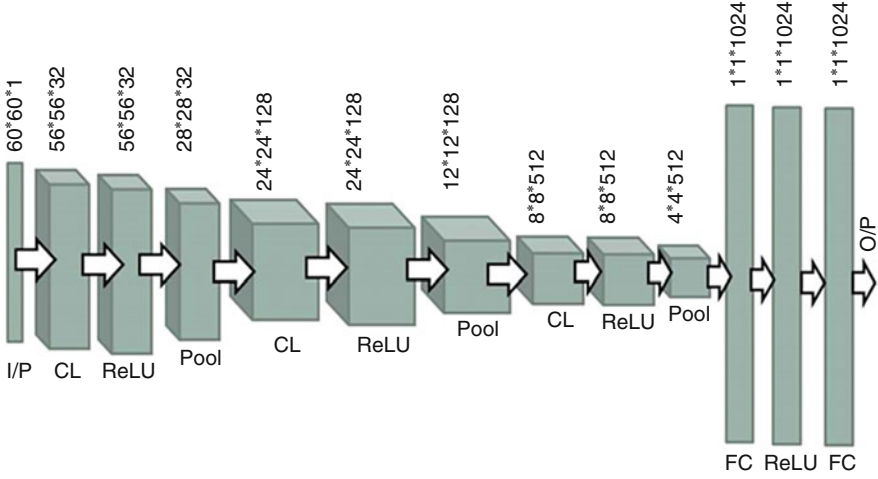


Fig. 7 The CNN topology

$$x_j^l = f(\beta_j^l \text{down}(x_i^{l-1}) + b_j^l) \quad (4)$$

where function $\text{down}(\cdot)$ denotes a max pooling function for our network. β_j^l is a weight and b_j^l is bias. The equation for max pooling function with a filter dimension $m \times m$ can be given by,

$$y = \max(x_i), i \in \{1, 2, \dots, m \times m\} \quad (5)$$

After max-pooling we get an output feature volume $28 \times 28 \times 32$.

Layer 4: The output of the pooling layer is fed to the 2nd convolution layer. With 128 filters of size 5×5 we get an output activation map $24 \times 24 \times 128$.

Layer 5: With the 2nd ReLU the nonlinear properties are further increased keeping the output volume same.

Layer 6: The 2nd max-pooling further reduces the number of features to an output volume $12 \times 12 \times 128$.

Layer 7: In the 3rd convolution layer the output of pooling layer is convolved with 512 filters of dimension 5×5 to get output activation $8 \times 8 \times 512$.

Layer 8: The ReLU changes the negative activation to 0 to further increase the nonlinearity.

Layer 9: The max-pooling down samples the input of ReLU to output volume 4×4 with receptive field size 2×2 .

Layer 10: This layer is a fully connected layer converted to a convolutional layer with a filter size 4×4 with 1024 kernels. It generates an activation map of size $1 \times 1 \times 1024$

Layer 11: The ReLU enhances the non-linear property.

Layer 12: The output of the ReLU is fed to another fully connected layer. From a single vector 1024 class scores, 2 classes: Normal and CRVO images are classified.

6 Result and Discussion

For our experiment, we collected 108 CRVO images (26 images from STARE database and 84 images from Retina Image Bank) and 133 Normal images (100 images from Hossein Rabbani database, 30 images from STARE database and 3 images from DRIVE database). We trained the CNN network with 100 normal (randomly selected from normal images from Hossein Rabbani, STARE and DRIVE databases) and 100 CRVO grayscale images (randomly selected from STARE and Retina Image Bank's CRVO images) of size 60×60 after preprocessing. In the 1st, 2nd and 3rd convolution layer, the filter size is 5×5 and in the 4th or last convolution layer/fully connected layer the filter size is 4×4 . The numbers of filters or kernels in the four convolution layers are 32, 128, 512 and 1024 respectively. The training is done with an epoch 70. For each training epoch we provided a batch size of nine training images and one validation image. Using the designed classifier we obtained an accuracy of 97.56%. Figure 8 shows the network training and validation for epoch 70. Figure 9 shows the Cumulative Match Curve (CMC) for rank vs. recognition rate. For the two classes we tested 41 images, 8 test images

Fig. 8 Network Training for epoch 70

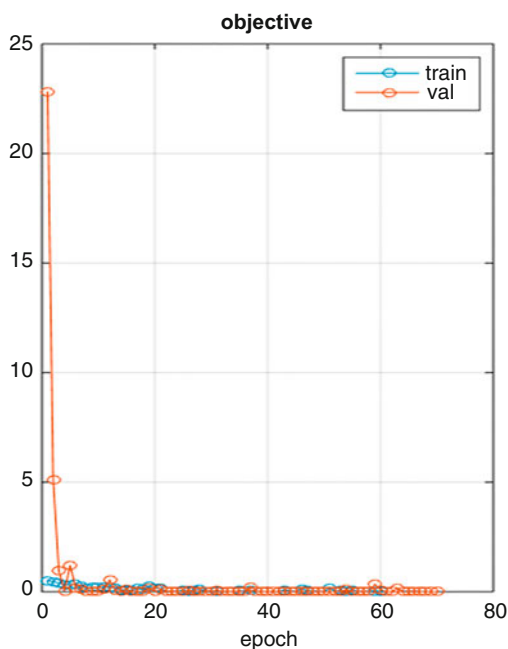
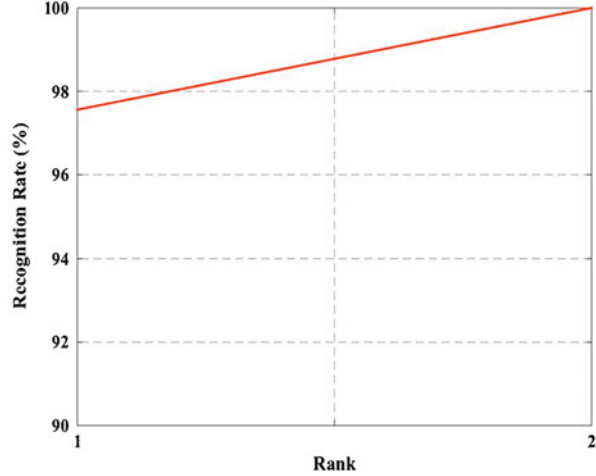


Fig. 9 Cumulative Match Curve (CMC)



for CRVO and 33 test images for normal retina. Each test produces a score for each image while comparing to each target class. If the score between test image and one of the target classes is larger than the other class, then that class is recognized in the first rank. Here, out of 41 test images 40 images are correctly recognized, hence the recognition rate for rank 1 is 97.56% and for rank 2 recognition rate is 100%. We further evaluated the performance in terms of specificity, sensitivity, positive predictive value and negative predictive value. Sensitivity is the probability of the positive test given that the patient has the disease. It measures the percentage of the people actually having the disease diagnosed correctly. Sensitivity can be given by following equation:

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (6)$$

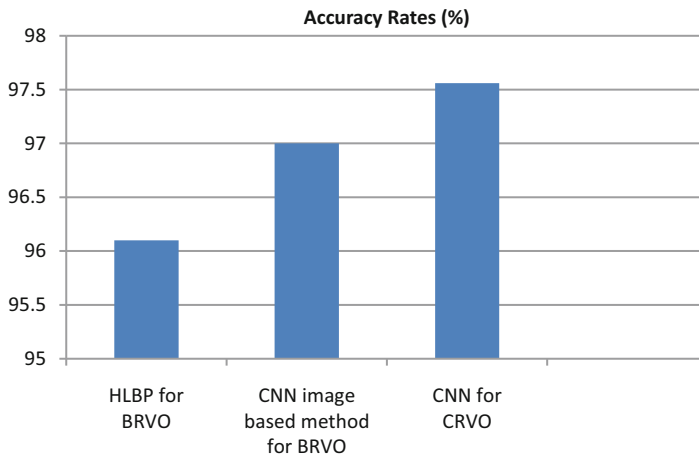
where, the “*True Positive*” depicts correctly identified disease and “*False Negative*” describes incorrectly rejected people having disease. In our experiment we got sensitivity 1. That means all the CRVO images are detected correctly. Again, specificity is the probability of a negative test given that the patient has no disease. It measures the percentage of the people not having disease diagnosed correctly. Specificity can be given by following equation:

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (7)$$

In our experiment, one normal image is incorrectly detected as CRVO image; hence, we obtained the specificity of 0.9697. The Positive Predictive value is the probability that subjects with a positive screening test truly have the disease. We got a positive predictive value 0.889. The Negative Predictive value is the probability

Table 2 Performance evaluation of the system

Accuracy	Sensitivity	Specificity	Error rate	Positive predictive value	Negative predictive value
97.56%	1	0.9697	2.44	0.8889	1

**Fig. 10** Accuracy rates of different methods for BRVO and proposed method for CRVO

that subjects with a negative screening test truly do not have the disease. We obtained negative predictive value 1. Table 2 summarizes the total evaluation of the system.

The experimental results show that the proposed method of detecting CRVO using CNN is a powerful method that we can implement in practice. Since there is no existing automatic detection of CRVO found in the literature, we are the first group to work on the automatic recognition of CRVO. Therefore, it is also difficult to compare the results with other methods. However, if we compare the method with that of automated recognition of BRVO, then our method performs better than the other feature extraction techniques and slightly better than the CNN based method. Figure 10 shows the comparison of our method with the existing methods for automated recognition of BRVO. So, the proposed method is fulfilling the need of automatic detection of CRVO to help the ophthalmologists in faster and efficient diagnosis of CRVO. It will also save the time and money of the patients. The method is taking care of the problems related to the image quality. This method is handling the most important issue, i.e., the requirement of different segmentation and feature extraction methods for detecting the abnormalities appear due to CRVO. Especially in the case, when detecting flame shaped haemorrhages, dilated veins and tortuous veins in the early stage of CRVO could be complicated and computationally expensive task. The supreme performance of the proposed CNN based method with a correct accuracy rate of 97.57% for the images of different sources proves it to be a promising consistent system for the automatic detection of CRVO. Because, all the images are captured by different funduscope devices and have different

types (image captured in different angles), format, resolution and quality. The performances of the existing automatic detection systems for BRVO are limited to a single dataset consisting of same type (image captured in same angle), same quality and same resolution images. Therefore, this CNN method is an efficient, versatile and consistent method for detecting CRVO.

7 Conclusion

In this chapter, we proposed a Central Retinal Vein Occlusion (CRVO) recognition method using Convolutional Neural Network (CNN). The designed network takes grayscale preprocessed images and recognizes the retina image with CRVO and the normal retina image. We have achieved a high accuracy of 97.56%. The proposed method is an image based method which is quite practical to implement. The advantage of this system is that there is no requirement of extra feature extraction step. The convolution layer serves both as feature extractor and the classifier. It is difficult to design feature extraction algorithm for the clinical signs of CRVO. Because, most of the time CRVO affects the whole retina and those large size hemorrhages, cotton wool spots are hard to define by other feature extraction methods. In CNN, each convolution layer extracts the low level features to the high level features from the CRVO images. Hence, it saves time. Since we conducted the experiment on retina image from different sources, the general automated detection method might affect the accuracy of the overall system due to different image quality, size and angle. However, use of CNN handles this situation due to its ability to cope with the distortions such as change in shape due to camera lens, different lighting conditions, different poses, presence of partial occlusions, horizontal and vertical shifts, etc. Therefore, the proposed CRVO detection scheme is a robust method.

References

1. The Royal College of Ophthalmologists: Retinal Vein Occlusion (RVO) Guidelines. The Royal College of Ophthalmologists, London (2015)
2. Sperduto, R.D., Hiller, R., Chew, E., Seigel, D., Blair, N., Burton, T.C., Farber, M.D., Gragoudas, E.S., Haller, J., Seddon, J.M., Yannuzzi, L.A.: Risk factors for hemiretinal vein occlusion: comparison with risk factors for central and branch retinal vein occlusion: the eye disease case-control study. *Ophthalmology*. **105**(5), 765–771 (1998)
3. Hayreh, S.S.: Prevalent misconceptions about acute retinal vascular occlusive disorders. *Prog. Retin. Eye Res.* **24**, 493–519 (2005)
4. Central Retinal Vein Occlusion Study Group: Natural history and clinical management of central retinal vein occlusion. *Arch. Ophthalmol.* **115**, 486–491 (1997)
5. Jelinek, H., Cree, M.J.: Automated image detection of retinal pathology. CRC Press, Boca Raton (2009)

6. Noma, H.: Clinical diagnosis in central retinal vein occlusion. *J. Med. Diagn. Methods.* **2**, 119 (2013)
7. Shiraishi, J., Li, Q., Appelbaum, D., Doi, K.: Computer-aided diagnosis and artificial intelligence in clinical imaging. *Semin. Nucl. Med.* **41**(6), 449–462 (2011)
8. Zhang, H., Chen, Z., Chi, Z., Fu, H.: Hierarchical local binary pattern for branch retinal vein occlusion recognition with fluorescein angiography images. *IEEE Electr. Lett.* **50**(25), 1902–1904 (2014)
9. Gayathri, R., Vijayan, R., Prakash, J.S., Chandran, N.S.: CLBP for retinal vascular occlusion detection. *Int. J. Comput. Sci. Iss.* **11**(2), 204–209 (2014)
10. Fazekas, Z., Hajdu, A., Lazar, I., Kovacs, G., Csakany, B., Calugaru, D.M., Shah, R., Adam, E.I., Talu, S.: Influence of using different segmentation methods on the fractal properties of the identified retinal vascular networks in healthy retinas and in retinas with vein occlusion. In: 10th Conference of the Hungarian Association for Image Processing and Pattern Recognition (KÉPAF 2015), pp. 360–373 (2015)
11. Zhao, R., Chen, Z., Chi, Z.: Convolutional neural networks for branch retinal vein occlusion recognition. In: IEEE International Conference on Information and Automation (2015)
12. Hijazi, S., Kumar, R., Rowen, C.: IP group cadence. https://ip.cadence.com/uploads/901/cnn_wp
13. LeNet details description: <http://deeplearning.net/tutorial/lenet.html>
14. Orth, D.H., Patz, A.: Retinal branch vein occlusion. *Surv. Ophthalmol.* **22**, 357–376 (1978)
15. A Beginner's Guide to Understanding Convolutional Neural Networks-Adit Deshpande of UCLA: <https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/>
16. Wu, Y., Hassner, T., Kim, K., Medioni, G., Natarajan, P.: Facial landmark detection with tweaked convolutional neural networks. *Comput. Vision Pattern Recogn.* **2**, preprint arXiv:1511.04031 (2015)
17. Ma, L., Minett, J.W., Blu, T., Wang, W.S: Resting state EEG-based biometrics for individual identification using convolutional neural networks. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2015)
18. Cecotti, H., Graser, A.: Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(3), 433–445 (2011)