COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 137 (2016) 281–292

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion

Pavle Prentašić *, *Sven Lončarić*

Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, 10000 Zagreb, Croatia

ARTICLE INFO

Article history:
Received 1 August 2016
Received in revised form
14 September 2016
Accepted 22 September 2016

Keywords:
Diabetic retinopathy
Exudates
Machine learning
Convolutional neural networks
Fundus photographs

ABSTRACT

Background and objective: Diabetic retinopathy is one of the leading disabling chronic diseases and one of the leading causes of preventable blindness in developed world. Early diagnosis of diabetic retinopathy enables timely treatment and in order to achieve it a major effort will have to be invested into automated population screening programs. Detection of exudates in color fundus photographs is very important for early diagnosis of diabetic retinopathy.

Methods: We use deep convolutional neural networks for exudate detection. In order to incorporate high level anatomical knowledge about potential exudate locations, output of the convolutional neural network is combined with the output of the optic disc detection and vessel detection procedures.

Results: In the validation step using a manually segmented image database we obtain a maximum F_1 measure of 0.78.

Conclusions: As manually segmenting and counting exudate areas is a tedious task, having a reliable automated output, such as automated segmentation using convolutional neural networks in combination with other landmark detectors, is an important step in creating automated screening programs for early detection of diabetic retinopathy.

1. Introduction

Diabetic retinopathy (DR) is a complication of diabetes caused by damage of small blood vessels in the retina. DR is one of the leading disabling chronic diseases and one of the leading causes of preventable blindness in the world [1]. In order to help reduce the burden for ever decreasing number specialists and achieve early diagnosis of diabetic retinopathy a major effort will have to be made in order to set-up screening programs for diabetic population. Yearly screening can be done by examining the eye fundus of the diabetic patients for any signs of diabetic retinopathy. It has been shown that fundus photography is the most accurate way of screening for diabetic retinopathy. Screening programs are important because almost one third of patients with diabetes may have DR changes without symptoms of vision problems [1]. DR is a progressive disease and because no symptoms of vision problems are

* *Corresponding author.* Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, 10000 Zagreb, Croatia. Fax: +38516129652.
E-mail address: pavle.prentasic@fer.hr (P. Prentašić).

http://dx.doi.org/10.1016/j.cmpb.2016.09.018
0169-2607/

Fig. 1 – Image with large number of exudates. In the image we can see many hard exudates, which appear as yellow spots with well defined border and soft exudates, which appear as pale yellow or white areas without well defined border. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

present efficient treatment is very difficult without early diagnosis. This is the main reason why a person with diabetes should have a comprehensive retina screening once in every year. Early diagnosis enables patients to maintain sufficient quality of vision and prevent severe vision loss and blindness [2], which eases the burden of the disease on patients and their families. Because the patient is still able to work and can live without other people's help, significant positive economical effects are achieved by early diabetic retinopathy detection and appropriate disease monitoring.

Diabetic retinopathy can progress from nonproliferative diabetic retinopathy (NPDR) to proliferative diabetic retinopathy (PDR). NPDR is the earliest stage of diabetic retinopathy, while PDR is an advanced stage of the disease. The DR starts with a mild NPDR, which does not affect the vision. As the disease progresses lipid and lipoprotein deposits called exudates appear near leaking capillaries within the retina. Exudates can be hard exudates, which appear as yellow spots with well defined border and soft exudates, which appear as pale yellow or white areas without well defined border. Their size is variable so the they can be as small as a few pixels but they can also be as large as the optic disc. Their shape and contrast levels can vary a lot. In Fig. 1 we can see how a typical image with exudates looks like.

We can find a lot of developed non-automatic screening programs for diabetic eye diseases in the literature [3–7]. The systems are implemented in different countries and for example in the UK, NHS Diabetic Screening Program offers annual fundus photography for all patients with diabetes over the age of 12, regardless of the socio-economic status [5]. In current screening programs only color fundus photography is used, and the data are sent to a grading center for reading where expert human readers estimate the disease severity. The main disadvantage is the necessity for qualified experts to grade the images, e.g. in the NHS Diabetes Screening Program patient's images can be graded by up to four different experts. This standard is impossible to achieve in countries with a shortage of qualified medical personnel.

The main contribution of this paper is a new method for exudate detection in color fundus photographs. The method uses convolutional neural networks and incorporates higher level knowledge about landmark points such as blood vessels and the optic disc in the detection procedure in order to increase the accuracy of exudate detection.

The paper is organized as follows: in Section 2 a short overview of different approaches for exudate detection is given. In Section 3 we show the details of our proposed method for exudate detection and explain how information about the landmark points is incorporated in the exudate detection procedure. In Section 4 we show the achieved results of the proposed method and we end the paper with a short conclusion.

2. Previous work

In the literature, we can find many different techniques for automatic detection of different pathologies in color fundus photographs, including detection of exudates. Most of the methods start with some sort of image preprocessing, followed by exudate candidate extraction procedure where a set of exudate candidates, i.e. structures, which are exudates or which are similar to exudates are extracted. Finally, a classification step is applied, where different features are extracted for each exudate candidate in order to keep the real exudate areas only.

The main goal of the preprocessing step is usually to reduce noise in input image but can also be used to increase the contrast of bright structures such as exudates. In color fundus photographs different bright structures are present such as the optic disc, which appears as a large, bright disc and is present in each image but also some structures such as drusen and optic nerve fibers, which do not appear in each image. Such bright structures can decrease the performance of exudate detection algorithms so one of the main goals of image preprocessing is to detect and remove such bright structures from original images. In Fig. 2 we can see several bright structures, such as drusen marked with arrow *b* and optic nerve fibers marked with arrow *a*, which come out of the optic disc and are usually visible near the main vessels. The optic disc is marked with a black circle.

In the literature we can find different preprocessing methods used in algorithms for detection of anatomic structure in retinal image of which many are presented in Ref. [8]. For example, authors use histogram equalization, adaptive histogram equalization, division by over-smoothed version of the original image, where the original intensity image is normalized by dividing the original intensity image by an over-smoothed version of the original image using a spatially large median filter [9], Gray World normalization [10], and illumination equalization where pixel values are modified by subtracting the mean intensity value calculated in a window.

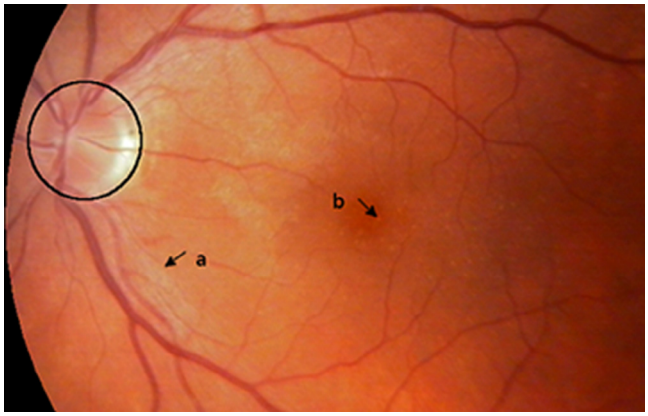

Fig. 2 – Bright structures in healthy patients: (a) Optic nerve fibers and (b) drusen.

After preprocessing, a candidate extraction procedure is applied in order to detect potential exudate areas. In the literature, we can find many different approaches for candidate extraction such as methods based on morphological operations [11–14], dynamic thresholding approaches [15], pixel-wise feature extraction and classification [13,16–18], clustering based approaches [19]. For example, in Ref. [11] authors start the exudate detection procedure by finding areas with standard deviation in a sliding window followed by a thresholding procedure. In order to find the precise border of each exudate area, morphological reconstruction by dilation is used where the generated binary image is used for morphological reconstruction. The reconstructed image is subtracted from the original image and after a final threshoding procedure a binary image containing exudate pixels is generated. In Ref. [15] authors used a mixture model for histogram modelling. They used the Expectation Maximization algorithm to estimate the mixture parameters. After parameter estimation, a dynamic threshold based on two Gaussian components with highest mixing weights is found and applied in order to segment the image into exudate and non-exudate regions. In Refs. [13,16] authors used pixel-wise features such as difference of Gaussians, standard deviation, hue, saturation, mean intensity values, maximum values, range of values in a window, entropy, edge strength and apply different classifiers such as K-Nearest Neighbours or Naive Bayes. In Ref. [19] authors applied the Fuzzy C-Means clustering algorithm on local contrast enhanced image. In Ref. [20] authors use the k-means clustering procedure on raw image pixels and Kirsch edge operator in order to remove areas with low edge strength.

After candidate extraction different features can be extracted for each connected component in order to eliminate spurious regions. Features such as area, length, perimeter, ratio of major and minor axis, average value inside the exudate area, mean and standard deviation of Gaussian derivative filter outputs in the candidate are extracted from each exudate candidate [21,22]. Some of the authors proposed features, which take into account the environment of the exudate candidate, for example distance from the blood vessel because exudates typically do not appear near the main vessels [23].

3. Materials and methods

We decided to combine outputs of different anatomical landmark detection algorithms with output of the exudate detection procedure in order to increase the accuracy of exudate detection. The general flowchart of the proposed method is presented in Fig. 3.

According to the flowchart we can see that we use the information about blood vessel locations, bright borders and optic disc location to increase the exudate detection performance. The exudate detection is done using a deep convolutional neural network, which is applied to the preprocessed image. The bright border probability map generation is explained in Subsection 3.3, parabola fitting in Subsection 3.4 and finally usage of convolutional neural networks is explained in Subsection 3.5.

Probability maps related to blood vessels, optic disc, parabola fitting and values for the bright border detection are combined together to get one probability map. We combine the probability map using a simple weighted sum of per pixel probability maps, where weights were found empirically. We have set the optic disc map weight to 0.3, blood vessel map weight to 0.3, parabola fitting mask weight to 0.2 and bright border detection map weight to 0.2. This map incorporates higher level semantic information about exudates. Example of such information is that exudates do not appear inside the optic disc or inside blood vessels. Also, exudates have lower probability of appearing near the main blood vessel.

Before starting the procedure for exudate detection we detect the optic disc, which usually appears as the largest and brightest circular structure in color fundus photographs. We segment the optic disc because the contrast and brightness of the optic disc can affect the performance of exudate detection method so it is a common practice to eliminate the optic disc area before applying the exudate detection algorithm. The proposed method is explained in more detail in Subsection 3.1.

We also segment the blood vessels because no exudates are present inside blood vessels. Bright structures such as the optic fibers appear near the main blood vessels. This information can be used to define regions with small probability of containing exudates. This probability map can improve the performance of exudate detection procedure by eliminating false positive detections. The proposed method for blood vessel segmentation is explained in Subsection 3.2.

3.1. *Optic disc detection*

In color fundus photographs the optic disc (OD) appears as a bright circle and in the literature we can find several optic disc detection algorithms [24–27]. Most of them try to find the optic disc based on color, shape, brightness or some other similar features. These algorithms work well when they are applied to healthy patients but they may fail when applied to images of low quality or images with a lot of visible pathologies or artifacts. In order to detect the optic disc we combine several optic disc detection algorithms into an ensemble. In the literature, such approaches have shown good results [28].

First we apply each optic disc detection algorithm separately and obtain optic disc probability maps for each algorithm applied. Then, we combine those probability maps into one final

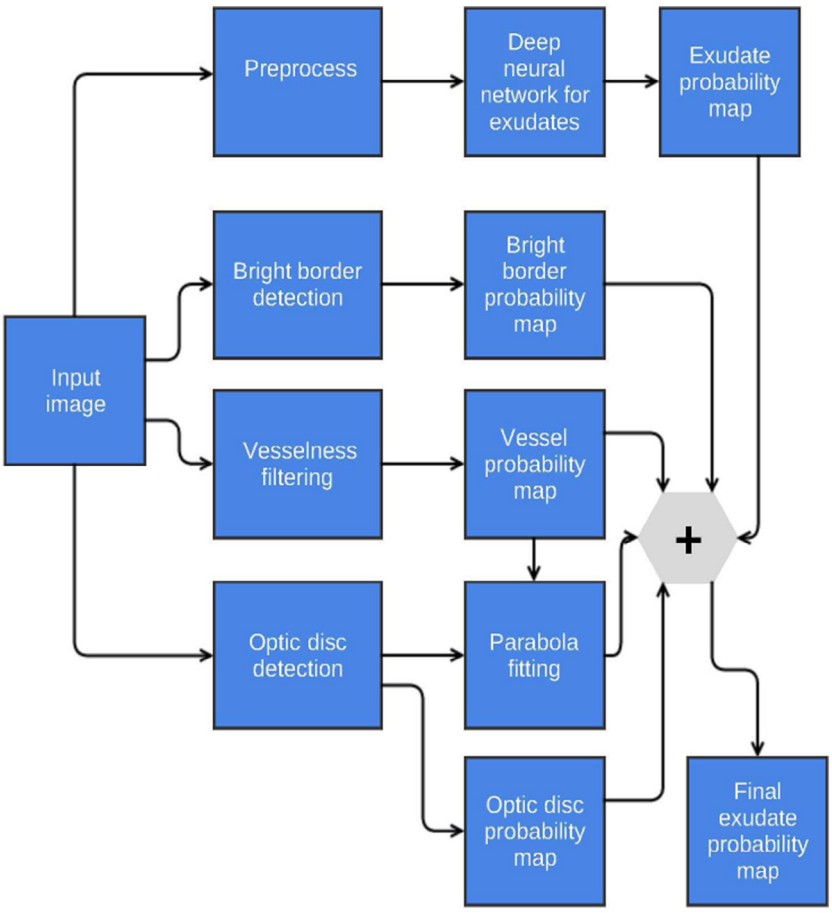

Fig. 3 – Flowchart of the proposed method. Outputs of vessel detection and optic disc detection methods are combined with output of the deep neural network in order to increase the accuracy of exudate detection.

probability map. The optic disc is located at the position corresponding to the maximum in the final probability map. The flowchart of the proposed method is shown in Fig. 4. In the following subsections we explain the methods, which are part of the ensemble.

3.1.1. *Entropy based approach*

This approach is based on entropy filtering proposed in Ref. [12]. The original RGB image is converted into Hue Saturation Intensity color space and after that median filtering is applied for noise removal. In order to increase local contrast, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to the intensity channel. After this step, entropy is calculated in a small sliding window. Regions with more details are represented with high entropy and smooth regions are represented with low entropy. In the paper, authors applied the Otsu's binarization algorithm in order to obtain a binary image. Because we want to preserve as much information as possible we just normalize the entropy values to 0–1 range to get the probability map.

3.1.2. *LoG filtering approach*

Because we assume that the optic disc is an object, approximately circular and consisting of bright pixels we can use generic methods for detection of blobs [29]. In our case, we apply

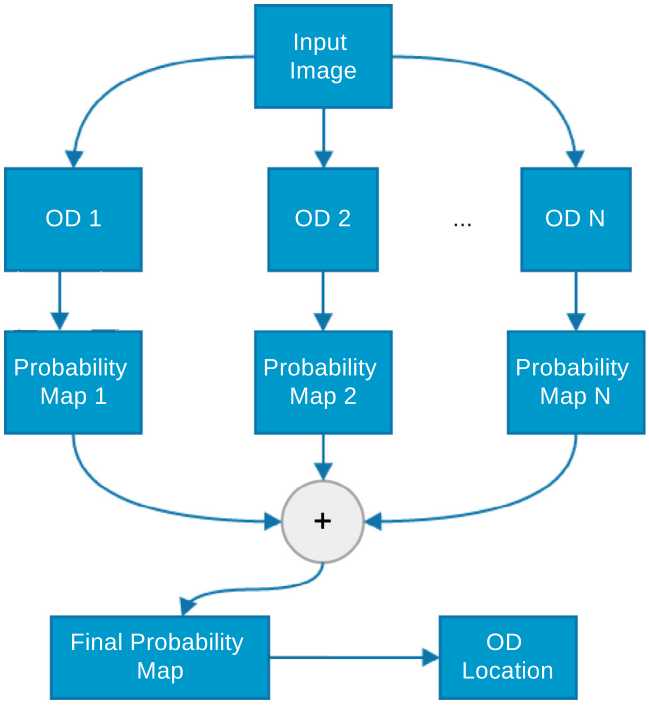

Fig. 4 – Flowchart of the proposed OD detection approach.

the Laplacian of Gaussian (LoG) filter to the green channel of the original image because this is the channel with the most local contrast. We use a circular convolution mask with radius similar to the expected radius of the optic disc because we want the mask to be similar to the optic disc. After we apply the filter, we just normalize the filtered image to 0–1 range to obtain the probability map.

3.1.3. *Brightness approach*

In Ref. [30] authors proposed an optic disc localization method, which applies a threshold to obtain pixels with high intensity and selects the center of the largest object as the optic disc center because the algorithm assumes that the optic disc is the largest bright part of the image. The detection is performed on the intensity component of the original image. In order to calculate the threshold, the histogram of the intensity channel is calculated and threshold is selected for which the thresholded image would contain the brightest two percent of pixels. The largest connected component is expected to be a part of the optic disc. In our case, we assign a value of one to pixels belonging to the largest connected object and zero to all other pixels in our probability map for this method.

3.1.4. *Hough transformation of vessels*

In Ref. [25] the authors used the Hough transformation of thinned blood vessels in order to detect the optic disc. This method assumes that the optic disc is the point of intersection of main blood vessels. In this paper authors detected blood vessels using a multi-scale based morphology approach. In our case, we used output of our blood vessel detection method as input to this method. After blood vessel thinning, Hough transform is performed and lines, which have a slope smaller than 45° are removed. For each line pair found after Hough transform intersection points are found and a vote is cast resulting in a voting map. The number of votes represents the number of lines crossing at this image point. In this approach, the highest value represent the center of the optic disc. In our case we normalize the voting map to 0–1 range to obtain a probability map.

3.1.5. *Probability map combining*

It is possible to combine information from probability maps in different ways. We decided to implement a straightforward method, which creates a new probability map by weighting different probability maps obtained by our optic disc detection methods. We decided to use the well known simulated annealing search algorithm [31] to find the the optimal weights for each probability map. We used the number of misclassified images as the energy function we were minimizing. The optimal weights found were 0.29 for the Entropy based approach, 0.16 for the LoG filtering approach, 0.14 for the Brightness approach and 0.41 for the Hough transformation of vessels approach.

The optic disc location is found by searching for the highest value in this new probability map or formally:

$$(x_{OD}, y_{OD}) = \arg\max_{(x,y)} [P(x, y) * d_{OD}] \quad (1)$$

where $P(x, y)$ is our final probability map and d_{OD} is a circular averaging filter, which we apply in order to smooth the final probability map. We use a circular averaging mask with radius similar to the expected radius of the optic disc because we want the mask to be similar to the optic disc. This center is used for parabola fitting. An example of the final optic disc probability map is visible in Fig. 5, while the corresponding input image is visible in Fig. 6.

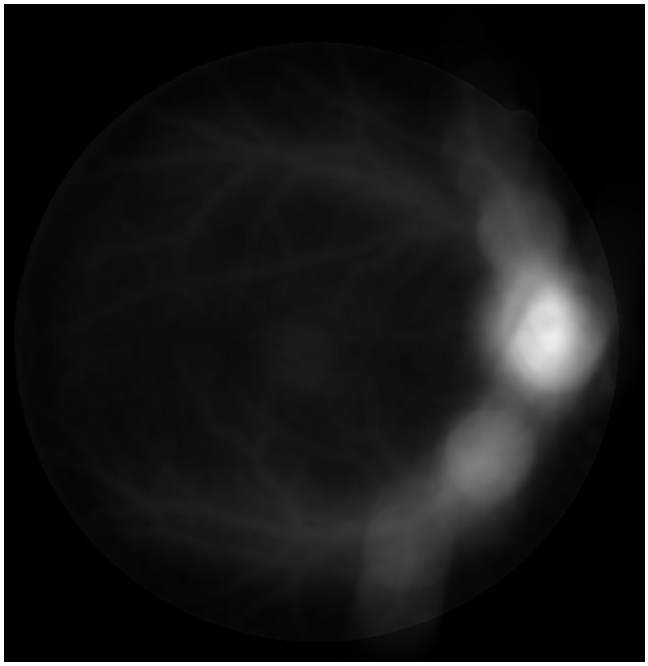

Fig. 5 – Optic disc probability map.

After the center of the optic disc has been found a circular mask is generated. The diameter of the mask was set to

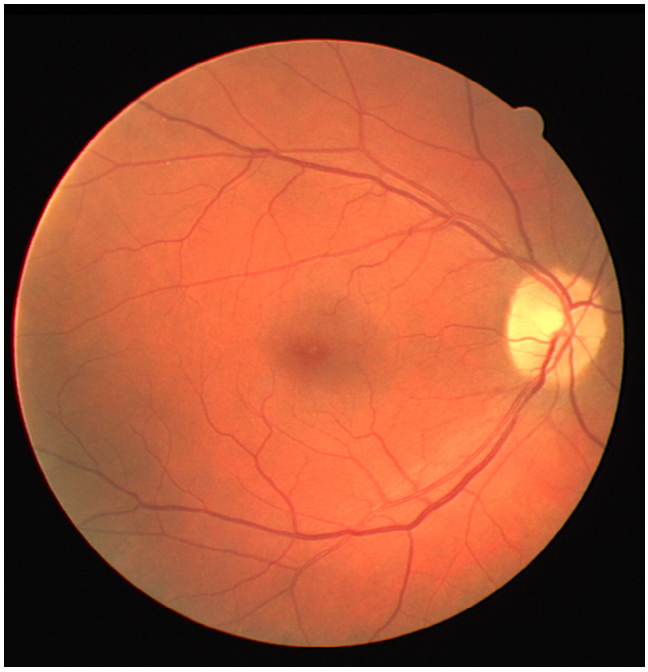

Fig. 6 – Example of a fundus image used for optic disc detection.

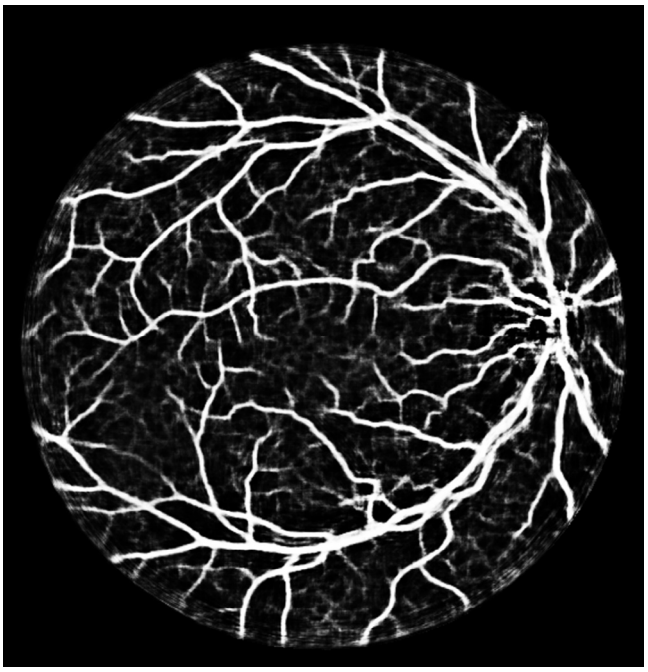

Fig. 7 – Output of the blood vessel detection procedure.

Fig. 8 – Example of a fundus image used for blood vessel detection.

$D/5$, where D is the diameter of the field of view as in Ref. [32]. This mask is later used for exudate detection because we assume that no exudate pixels can appear inside of the optic disc. Values inside of the mask are set to zero and points outside of the mask are set to one. We can think about this mask as a probability mask where each point indicates the probability of that point lying inside of the optic disc.

3.2. *Detection of blood vessels*

Blood vessels are one the main landmarks in blood vessels and are darker in appearance to background. Segmentation of blood vessels in our case is important for optic disc localization, parabola fitting and generating adaptive exclusion masks because exudates do not appear inside blood vessels. Because we are mostly interested in segmentation of main vessel we could have used one of many algorithms presented in the literature [33–35] but we decided to use Frangi vesselness filter in order to increase the contrast level of blood vessels [36] due to its simplicity and speed. Example of the output of the Frangi vesselness filter applied to the fundus image can be seen in Fig. 7. These outputs can be treated as probabilities. Higher values mean higher probability of a pixel being part of a blood vessel. The original image is visible in Fig. 8.

3.3. *Detection of bright borders*

Some images have extremely bright regions along the border of the visible part of the fundus image. Those regions can cause problems with exudate detection because they are very bright. Example of such a border is marked in Fig. 9. In order to account for such anomalies we segment the bright border if it is present in the image. This segmentation is then used as part of the exudate detection procedure.

We build upon the method presented in Ref. [32] so in our approach we take the green channel of the original image and apply a large median filter to the green channel in order to estimate the background of the image. After that we subtract the estimated background from the green channel in order to remove the background from the image. We keep only positive values and set all other values to zero. Because we are only interested in bright areas around edges of field of view we need to eliminate the bright regions inside of the retina. We do that by creating an attenuation map for each pixel of our image. Values of the proposed map are calculated using Eq. (2) where $S(x, y)$ is the value of the attenuation at point (x, y), $D(x, y)$ is the distance from the closest point to the edge of the field of

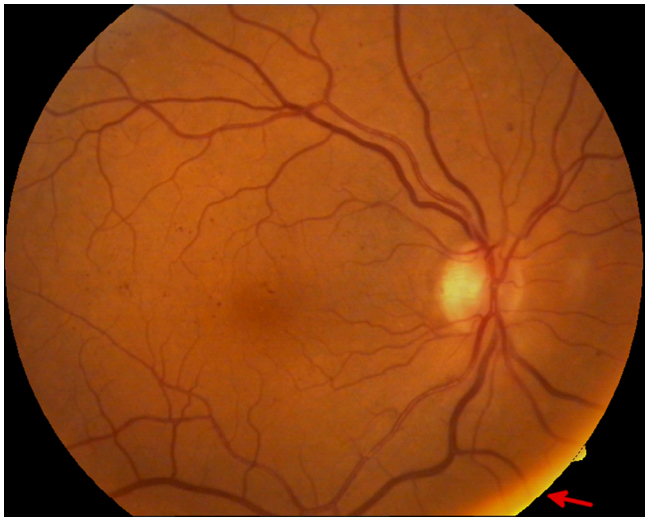

Fig. 9 – Image with bright border marked.

Fig. 10 – Bright border segmented.

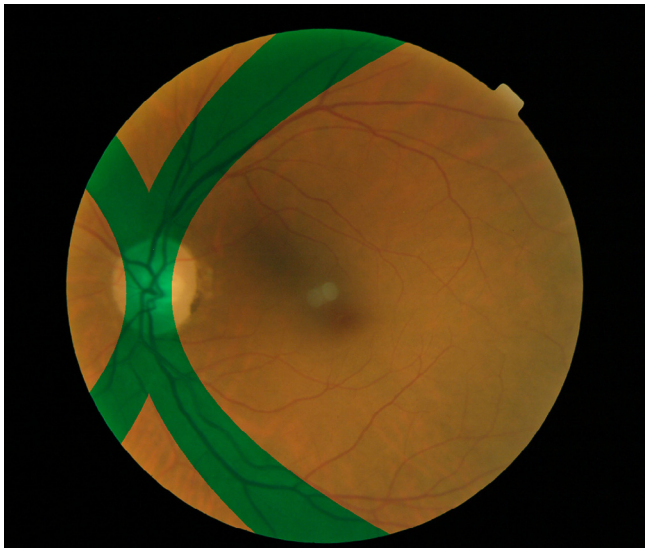

Fig. 11 – Parabola fitted to main vessels.

view and a is a positive constant. In our experiments, a was set to 0.02.

The field of view mask can be easily calculated using simple thresholding. Image pixels outside of the field of view are zero or close to zero so applying a low threshold will segment the field of view. Our mask is just the largest blob in the obtained binary image. We apply the fill operation inside of the blob to fill any potential holes in the mask.

$$S(x, y) = \exp^{-a \cdot D(x,Y)} \tag{2}$$

After that, we just apply the attenuation map to the green channel with estimated background subtracted, which creates a new image in which only bright regions near the edge of the field of view are visible. In order to preserve only bright pixels we apply a fixed threshold to this new image. Pixels with values less than the threshold are set to zero and all other pixels are left unchanged. The threshold was empirically found. In this new image, higher values represent bright points near the border. Example of such an image can be seen in Fig. 10.

3.4. *Parabola fitting*

Most reflections and optic fibers are found along the main retina vessels and they lie inside a parabolic region passing through the optic disc. After the optic center is localized and blood vessels detected we fit a symmetric double parabolic region passing through the optic disc center. We fit the parabola to largest blood vessels, which converge around the optic disc area, which is similar to the approach presented in Ref. [37].

The parabola fitting algorithm starts with the segmented blood vessels. In order to segment the blood vessels we start with the vessel probability map. This map is thresholded using an adaptive threshold. We choose a threshold such that 5% of the image is covered by blood vessels. This is done by sorting the probability map in a descending order and selecting the threshold such that 5% of values in the sorted map are larger than the selected threshold. Because we are only interested in main blood vessels we only take 10% of thickest blood vessels. This is done by iteratively eroding the binary vessel image. After this step we thin the blood vessels so all blood vessels are only one pixel thin. The remaining points after thinning are the points used to estimate the parabola. Because our parabola is shifted to the optic disc center (x_{OD}, y_{OD}) the resulting parabola equation is given by Eq. (3).

$$a \cdot (y - y_{OD})^2 = |x - x_{OD}| \tag{3}$$

The only parameter to be minimized is the scaling factor a. We find this parameter by applying the standard nonlinear Marquardt [38] optimization procedure to the criterion function J from Eq. (4) where S is a set of points in blood vessel image after thinning. The optimization procedure converges quickly, usually in less than 10 iterations. In Fig. 11 we can see a parabolic region drawn on top of the original image.

$$J(a) = \sum_{(x,y) \in S} a \cdot (y - y_{OD})^2 - |x - x_{OD}| \tag{4}$$

The main purpose of parabola fitting is to define regions with low exudate probability. The regions around main blood vessels tend to contain many reflections and optic fibers, which can decrease the performance of the exudate detection algorithm. In our case we create an adaptive probability mask for those regions based on parabola fitting of main vessels. So regions, which are close to the parabola should have low probability of exudate appearance, and regions far away should have high probability of exudate appearance. We model this by using a distance function from the fitted parabola. We model the probability by using the distance function in Eq. (5).

$$P(x, y) = 1 - \exp^{-a \cdot D(x,y)} \tag{5}$$

Here, $P(x, y)$ is the a priori probability of point (x, y) containing exudates, $D(x, y)$ is the Euclidean distance of point (x, y) from the parabola fitted to the main vessels and a is a small positive constant. Example of the proposed a priori

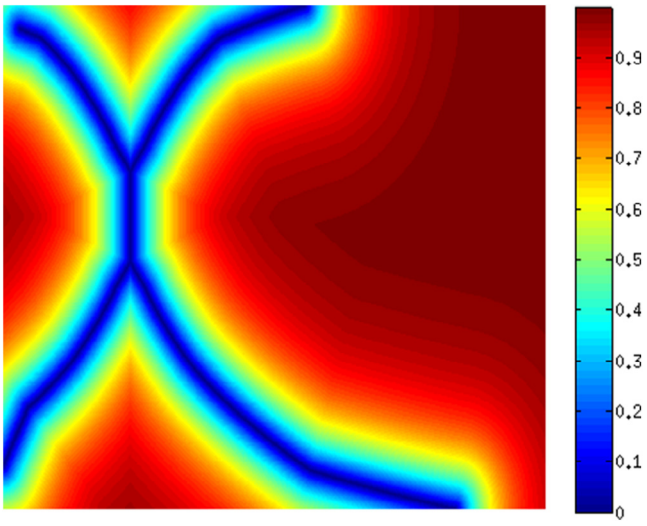

Fig. 12 – A priori exudate probability function.

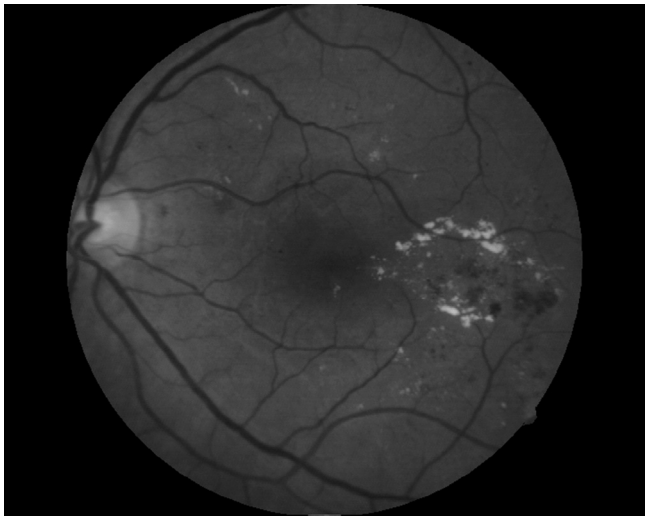

Fig. 13 – Original noisy image.

probability function can be seen in Fig. 12. Blue values mean that there is a small probability of finding exudates in that area and red values mean that there is a high probability of finding exudates in that area. This probability map is later combined with output of the exudate detection procedure to increase the accuracy of exudate detection.

We have chosen the exponential function as the generator for our a priori probability map because the exponential function has only one tunable parameter (small positive constant a), which was empirically found.

3.5. *Exudate detection procedure*

We decided to use convolutional neural networks for exudate detection in color fundus photographs where feature extraction is learned from data and not enforced by designers. These approaches obtained state of the art results in a very broad range of applications [39]. Before applying the convolutional neural network we preprocess the original input image.

3.5.1. *Preprocessing*

In order to reduce the noise levels in fundus photographs before using the convolutional neural networks for exudate detection we use the Total Variation (TV) regularization denoising, which was originally developed for additive white Gaussian noise denoising by Rudin, Osher and Fatemi [40]. The authors proposed to estimate the denoised image u as the solution to the minimization problem:

$$\arg\min_{u\in BV(\Omega)} \|u\|_{TV(\Omega)} + \frac{\lambda}{2}\int_{\Omega}(f(x)-u(x))^2\,dx \tag{6}$$

where λ is a positive parameter. Here, f is the observed noisy image, which is related to the underlying true image u by $f = u + \eta$, and η is at each point in space independently and identically distributed as a zero-mean Gaussian random variable. This problem is referred to as the Rudin–Osher–Fatemi or ROF problem. In our case total variation is the integral of its gradient magnitude:

$$\|u\|_{TV(\Omega)} = \int_{\Omega}|\nabla u|dx \tag{7}$$

Such formulation of the minimization function discourages the solution from having oscillations, yet it does allow the solution to have discontinuities. This is possible because if u is monotonic in $[a, b]$ then $TV(u) = |u(b) - u(a)|$, regardless of whether u is discontinuous or not. The second term in Eq. (6) encourages the solution to be close to the observed image f. By this combination, the minimization finds the denoised image.

In order to perform the denosing the split Bregman algorithm is used, which is explained in more detail in Refs. [41,42].

In Fig. 13 we can see how the original image looks like, and in Fig. 14 we can see how the denoised image looks like. Small section of the original and denoised image is shown in Fig. 15. Obviously, we can see that noise levels are reduced and all edges are still present. This is important because hard edges are one of the main characteristics of exudates.

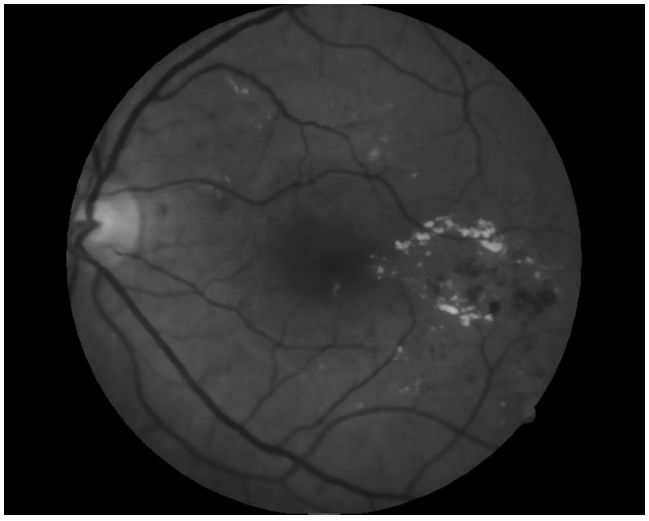

Fig. 14 – Denoised image.

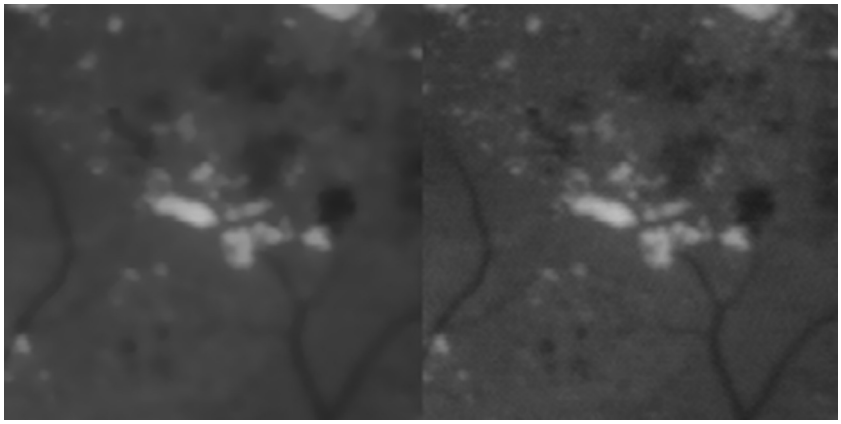

Fig. 15 – Denoised image (left) and noisy image (right).

After preprocessing, we apply a deep neural network, or to be more specific we use a convolutional neural network (CNN) for exudate detection. Our method is inspired by work presented in Ref. [43] where authors use deep neural networks in order to segment neuronal membranes in electron microscopy. So, we want to classify each pixel in *exudate* or *non-exudate* class. Our convolutional neural network calculates the probability of a pixel being one of the mentioned two classes. Inputs of our network are raw intensity values in the total variation preprocessed image of a square window centered in the pixel p, which we are currently processing. The size of the window is an odd number in order to enforce symmetry around the given pixel. For pixels near the image border, our window would include pixels outside the image boundaries. In those cases we insert new values in our sliding window by mirroring the pixels from the window which falls inside the image.

The convolutional neural network is first trained using several images from the training set. After training, in order to segment the image, we have to apply our convolutional neural network for each pixel. This means that output of the image classification is another image where each pixel value represents the probability of pixel being an exudate. In order to get a binary image a fixed threshold would have to be applied. In our case we use the the probability maps for blood vessels, optic disc, parabola fitting and bright border as a priori information, which is combined with the output of the convolutional neural network in order to create the final exudate probability map. This map can then be thresholded in order to get the exudate areas.

3.5.2. *Convolutional neural network architecture*

A typical convolutional neural network consists of a sequence of convolutional, max-pooling and fully connected layers [43]. This type of deep neural network is hierarchical feature extractor, which uses raw pixel intensities of the original image in order to create a new feature vector, which is then classified by several fully connected layers. This is the main difference compared to other machine learning approaches where we have to manually decide on the features to be used in a classifier because, in the convolutional neural networks the filters are automatically learned from training data. Each convolutional layer from the network performs a 2D convolution of its input images called input maps with a sequence of square filters. Output of each input map is calculated by summing the convolutional responses over the whole input map. After that, this sum is passed through a non-linear activation function. Finally, a max-pooling layer downsamples the output of non-linear activation function by a constant factor. Their outputs are given by the maximum activation over non-overlapping square regions. Max-pooling layers are fixed, non-trainable layers, which select the most promising features [44]. After a few stages of alternating convolutional and max-pooling layers outputs of the final max-pooling layer are brought as inputs to a sequence of fully connected layers. The output layer is a fully connected layer with one neuron per class, which in our case is two neurons. We add a softmax activation function after the last layer so that the output of our convolutional neural network can be treated as the probability of a particular image pixels belonging to the exudate class. The architecture of our convolutional neural network is visible in Table 1. We can see that there are four convolutional layers and four max-pooling layers. The size of the input map is 65×65, which means that in order to classify a pixel, convolution in a window of size 65×65 is performed in the input layer. As it was mentioned, before max-pooling we apply a non-liner activation function, which in our case is a rectifying linear unit.

To train our convolutional neural network, we use all available positive training samples from our training images. From

Table 1 – Network layers architecture.

Layer	Type	Maps and size	Kernel size
0	Input	1 map of 65×65 neurons	–
1	Convolutional	48 maps of 60×60 neurons	6×6
2	Max pooling	48 maps of 30×30 neurons	2×2
3	Convolutional	48 maps of 26×26 neurons	5×5
4	Max pooling	48 maps of 13×13 neurons	2×2
5	Convolutional	48 map of 10×10 neurons	4×4
6	Max pooling	48 maps of 5×5 neurons	2×2
7	Convolutional	48 maps of 4×4 neurons	2×2
8	Max pooling	48 maps of 2×2 neurons	2×2
9	Fully connected	100 neurons	–
10	Fully connected	2 neurons	–

each image, we take all the exudate pixels as positive samples and the same amount of pixels randomly sampled among all non-exudate pixels but without repetition. We perform this in order to have a balanced training set. Because exudates can appear in different shapes and orientation in order to augment the training set we synthetically add rotated and mirrored versions of training samples. Positive and negative samples are interleaved to have approximately equal number of positive and negative samples when randomly sampling the training set. We take the green channel as the input to our network because it contains the most contrast according to the literature [45].

Initial convolutional kernel weights were drawn from a Gaussian distribution with zero mean and standard deviation of 0.01 and initial weights of fully connected layers were drawn from a Gaussian distribution with zero mean and standard deviation of 0.1. Batch size during the training phase was set to 100 and the network was trained for 200,000 iterations. The training was stopped when there was no significant improvement of accuracy on the validation set.

4. Results

Training and testing of the proposed method was done using a computer with a Tesla K20C graphics card in order to speed the computation. We used the Caffe deep learning toolkit [46] in order to efficiently use the processing power of the Tesla graphics card for computation of convolutional neural network parameters. It takes approximately 10 hours in order to train the proposed neural network using the mentioned hardware.

We tested our method on the DRiDB database [47], which is available on request. This dataset contains 50 color fundus images for which all the main structures like blood vessels, optic disk and macula are marked along with pathological changes like hard and soft exudates, dot and blot hemorrhages and neovascularizations. We split the database into two disjoint sets for training and testing purposes.

To test our method we use the ground truth data available in the mentioned datasets. For each image, the number of true positives (TP), false positives (FP) and false negatives (FN) is calculated. Measuring the true positives, false positives and false negative could be done by counting the number of pixels, which are correctly classified but this approach has some drawbacks because it can happen that we actually detect the exudate blob correctly but because the ground truth segmentation is not perfect several border pixels can be assigned as false positive or false negative. In order to solve this problem we use the approach explained in Ref. [32]. Each segmented image can be divided into a set of candidates $\{C_1, C_2, \ldots, C_N\}$ where each C_i represents a connected component. Each ground truth image can also be divided into a set of candidates $\{G_1, G_2, \ldots, G_M\}$. So our ground truth mask is given by:

$$G = \bigcup_{1 \le j \le M} G_j \tag{8}$$

Our segmented images can be represented as:

Table 2 – Results of different exudate detection methods.

Method name	Sensitivity	PPV	F-score
Walter [11]	0.69	0.48	0.57
Sánchez [48]	0.34	0.61	0.44
Harangi [13]	0.66	0.65	0.66
Harangi [22]	0.71	0.66	0.68
Amel [20]	0.41	0.09	0.15
Proposed method	0.78	0.78	0.78

$$C = \bigcup_{1 \le i \le N} C_i \tag{9}$$

We consider a pixel to be a true positive pixel if, and only if, it belongs to any of the following sets:

- $C \cap G$
- C_i such that $\frac{|C_i \cap G|}{|C_i|} > \sigma$
- G_j such that $\frac{|G_j \cap C}{G_j} > \sigma$

Here, |,| represents number of elements in set and σ is a parameter in $[0, 1]$ range. We consider a pixel to be a false positive pixel if, and only if, it belongs to any of the following sets:

- C_i such that $C_i \cap G = \varnothing$
- $C_i \cap \bar{G} \le \sigma$

We consider a pixel to be false negative if, and only if, it belongs to any of the following sets:

- G_j such that $G_j \cap C = \varnothing$
- $G_j \cap \bar{C}$ such that $\frac{|G_j \cap C|}{|G_j|} \le \sigma$

We omit the true negative from our analysis because the number of true negatives can be very high since all non exudate pixels are actually true negatives. We set σ to 0.2.

We compute the sensitivity S of detection as:

$$S = \frac{TP}{TP + FN} \tag{10}$$

and the positive predictive value (PPV) as:

$$PPV = \frac{TP}{TP + FP} \tag{11}$$

Finally we calculate the F-score as:

$$F = \frac{2 \cdot S \cdot PPV}{PPV + S} \tag{12}$$

Table 2 presents the results of the experimental validation after 3-fold cross-validation for the DRiDB database. The proposed method outperforms all algorithms used in the validation process. In Fig. 16 we can see the results of the proposed method.

Fig. 16 – Result obtained using the proposed method. Green: True positives, Blue: False positives, Cyan: False negatives. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusion

The segmentation of exudates in fundus photographs has been a heavily researched area in recent years. Although many techniques and algorithms have been developed, there is still room for further improvements.

In this paper we showed that deep convolutional neural networks can be effectively used in order to segment exudates in color fundus photographs. In the future we will try to enhance our network by using all the available channels and adding some preprocessing and postprocessing steps because grouping detected pixels in clusters and adding some high level features would certainly improve the final segmentation result.

REFERENCES

[1] S. Wild, G. Roglic, A. Green, R. Sicree, H. King, Global prevalence of diabetes: estimates for the year 2000 and projections for 2030, Diabetes Care 27 (5) (2004) 1047–1053.

[2] C.A. McCarty, C.W. Lloyd-Smith, S.E. Lee, P.M. Livingston, Y.L. Stanislavsky, H.R. Taylor, Use of eye care services by people with diabetes: the Melbourne Visual Impairment Project, Br. J. Ophthalmol. 82 (4) (1998) 410–414.

[3] D.A. Askew, L. Crossland, R.S. Ware, S. Begg, P. Cranstoun, P. Mitchell, et al., Diabetic retinopathy screening and monitoring of early stage disease in general practice: design and methods, Contemp. Clin. Trials 33 (5) (2012) 969–975.

[4] H.C. Looker, S.O. Nyangoma, D. Cromie, J.A. Olson, G.P. Leese, M. Black, et al., Diabetic retinopathy at diagnosis of type 2 diabetes in Scotland, Diabetologia 55 (9) (2012) 2335–2342.

[5] T. Peto, C. Tadros, Screening for diabetic retinopathy and diabetic macular edema in the United Kingdom, Curr. Diab. Rep. 12 (4) (2012) 338–345.

[6] M.D. Abràmoff, M. Niemeijer, M.S. Suttorp-Schulten, M.A. Viergever, S.R. Russell, B. Van Ginneken, Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes, Diabetes Care 31 (2) (2008) 193–198.

[7] G.S. Scotland, P. McNamee, S. Philip, A.D. Fleming, K.A. Goatman, G.J. Prescott, et al., Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in Scotland, Br. J. Ophthalmol. 91 (11) (1518–1523) 2007.

[8] A.A. Youssif, A.Z. Ghalwash, A.S. Ghoneim, Comparative study of contrast enhancement and illumination equalization methods for retinal vasculature segmentation, in: Proceedings of the Third Cairo International Biomedical Engineering Conference (CIBEC'06), 2006, pp. 1–5.

[9] Z. Xiao, F. Li, L. Geng, F. Zhang, J. Wu, X. Zhang, et al., Hard exudates detection method based on background-estimation, in: International Conference on Image and Graphics, Springer, 2015, pp. 361–372.

[10] K.A. Goatman, A.D. Whitwam, A. Manivannan, J.A. Olson, P.F. Sharp, Colour normalisation of retinal images, in: Proceedings of Medical Image Understanding and Analysis, 2003, pp. 49–52.

[11] T. Walter, J.-C. Klein, P. Massin, A. Erginay, A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina, IEEE Trans. Med. Imaging 21 (10) (2002) 1236–1243.

[12] A. Sopharak, B. Uyyanonvara, S. Barman, T.H. Williamson, Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods, Comput. Med. Imaging Graph. 32 (8) (2008) 720–727.

[13] B. Harangi, B. Antal, A. Hajdu, Automatic exudate detection with improved Naïve-Bayes classifier, in: Computer-Based Medical Systems (CBMS), 2012 25th International Symposium, IEEE, 2012, pp. 1–4.

[14] R. Annunziata, A. Garzelli, L. Ballerini, A. Mecocci, E. Trucco, Leveraging multiscale hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation, J. Biomed. Health Inform. 20 (2015) 1129–1138.

[15] C.I. Sánchez, M. García, A. Mayo, M.I. López, R. Hornero, Retinal image analysis based on mixture models to detect hard exudates, Med. Image Anal. 13 (4) (2009) 650–658.

[16] M. Niemeijer, B. van Ginneken, S.R. Russell, M.S. Suttorp-Schulten, M.D. Abràmoff, Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis, Invest. Ophthalmol. Vis. Sci. 48 (5) (2007) 2260–2267.

[17] A. Sopharak, B. Uyyanonvara, S. Barman, Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy c-means clustering, Sensors (Basel) 9 (3) (2009) 2148–2161.

[18] P.M. Rokade, R.R. Manza, Automatic detection of hard exudates in retinal images using Haar wavelet transform, Int. J. Appl. Innov. Eng. Manage. 4 (5) (2015) 402–410.

[19] A. Osareh, M. Mirmehdi, B. Thomas, R. Markham, Automatic recognition of exudative maculopathy using fuzzy c-means clustering and neural networks, in: Proceedings of Medical Image Understanding Analysis Conference, vol. 3, 2001, pp. 49–52.

[20] F. Amel, M. Mohammed, B. Abdelhafid, Improvement of the hard exudates detection method used for computer-aided diagnosis of diabetic retinopathy, Int. J. Image Graph. Signal Process. 4 (4) (2012).

[21] C.I. Saánchez, M. Niemeijer, M.S. Suttorp Schulten, M. Abraámoff, B. van Ginneken, Improving hard exudate

detection in retinal images through a combination of local and contextual information, in: Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium, IEEE, 2010, pp. 5–8.
[22] B. Harangi, I. Lazar, A. Hajdu, Automatic exudate detection using active contour model and regionwise classification, in: Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, IEEE, 2012, pp. 5951–5954.
[23] C.I. Sánchez, M. Niemeijer, I. Išgum, A. Dumitrescu, M.S. Suttorp-Schulten, M.D. Abràmoff, et al., Contextual computer-aided detection: improving bright lesion detection in retinal images and coronary calcification identification in ct scans, Med. Image Anal. 16 (1) (2012) 50–62.
[24] A. Hoover, M. Goldbaum, Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels, IEEE Trans. Med. Imaging 22 (8) (2003) 951–958.
[25] S. Ravishankar, A. Jain, A. Mittal, Automated feature extraction for early detection of diabetic retinopathy in fundus images, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009, 2009, pp. 210–217.
[26] Y. Abdel-Razik, A.Z. Ghalwash, G. Abdel-Rahman, Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter, IEEE Trans. Med. Imaging 27 (1) (2008) 11–18.
[27] A. Osareh, M. Mirmehdi, B. Thomas, R. Markham, Comparison of colour spaces for optic disc localisation in retinal images, in: 16th International Conference on Pattern Recognition, 2002. Proceedings, vol. 1, 2002, pp. 743–746.
[28] B. Harangi, A. Hajdu, Detection of the optic disc in fundus images by combining probability models, Comput. Biol. Med. 65 (2015) 10–24.
[29] R.M. Haralick, L.G. Shapiro, Computer and Robot Vision, vol. 1, Addison-Wesley Longman Publishing Co., Inc., 1992.
[30] T. Walter, J.-C. Klein, Segmentation of color fundus images of the human retina: detection of the optic disc and the vascular tree using morphological techniques, in: Medical Data Analysis, Springer, 2001, pp. 282–287.
[31] W.L. Goffe, G.D. Ferrier, J. Rogers, Global optimization of statistical functions with simulated annealing, J. Econom. 60 (1) (1994) 65–99.
[32] X. Zhang, G. Thibault, E. Decenciere, B. Marcotegui, B. Laÿ, R. Danno, et al., Exudate detection in color retinal images for mass screening of diabetic retinopathy, Med. Image Anal. 18 (7) (2014) 1026–1043.
[33] U.T. Nguyen, A. Bhuiyan, L.A. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, Pattern Recognit. 46 (3) (2013) 703–715.
[34] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, M.D. Abramoff, Comparative study of retinal vessel segmentation methods on a new publicly available database, in: Medical Imaging 2004, International Society for Optics and Photonics, 2004, pp. 648–656.
[35] E. Bekkers, R. Duits, T. Berendschot, B. ter Haar Romeny, A multi-orientation analysis approach to retinal vessel tracking, J. Math. Imaging Vis. 49 (3) (2014) 583–610.
[36] A.F. Frangi, W.J. Niessen, K.L. Vincken, M.A. Viergever, Multiscale vessel enhancement filtering, in: Medical Image Computing and Computer-Assisted Interventation—MICCAI'98, Springer, 1998, pp. 130–137.
[37] K.W. Tobin, E. Chaum, V.P. Govindasamy, T.P. Karnowski, Detection of anatomic structures in human retinal imagery, IEEE Trans. Med. Imaging 26 (12) (2007) 1729–1739.
[38] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Indust. Appl. Math. 11 (2) (1963) 431–441.
[39] J. Masci, A. Giusti, D. Cireşan, G. Fricout, J. Schmidhuber, A fast learning algorithm for image segmentation with max-pooling convolutional networks, arXiv preprint arXiv:1302.1690, 2013.
[40] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1) (1992) 259–268.
[41] T. Goldstein, S. Osher, The split Bregman method for l1-regularized problems, SIAM J. Imaging Sci. 2 (2) (2009) 323–343.
[42] P. Getreuer, Rudin-Osher-Fatemi total variation denoising using split Bregman, Image Process. Online 10 (2012) 74–95. http://dx.doi.org/10.5201/ipol.2012.g-tvd.
[43] D. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Deep neural networks segment neuronal membranes in electron microscopy images, in: Advances in Neural Information Processing Systems, 2012, pp. 2843–2851.
[44] D. Ciresan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference, IEEE, 2012, pp. 3642–3649.
[45] R.J. Winder, P.J. Morrow, I.N. McRitchie, J. Bailie, P.M. Hart, Algorithms for digital image processing in diabetic retinopathy, Comput. Med. Imaging Graph. 33 (8) (2009) 608–622.
[46] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, et al., Caffe: convolutional architecture for fast feature embedding, arXiv preprint arXiv:1408.5093, 2014.
[47] P. Prentasic, S. Loncaric, Z. Vatavuk, G. Bencic, M. Subasic, T. Petkovic, et al., Diabetic retinopathy image database (DRiDB): a new database for diabetic retinopathy screening programs research, in: Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium, IEEE, 2013, pp. 711–716.
[48] C.I. Sánchez, R. Hornero, M.I. López, M. Aboy, J. Poza, D. Abasolo, A novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis, Med. Eng. Phys. 30 (3) (2008) 350–357.