An Automated System for the Detection and Classification of Retinal Changes Due to Red Lesions in Longitudinal Fundus Images

Adal, Kedir M.; Van Etten, Peter G.; Martinez, Jose P; Rouwen, Kenneth W.; Vermeer, Koenraad A.; van Vliet, Lucas J.

DOI
10.1109/TBME.2017.2752701

Publication date
2017

Document Version
Accepted author manuscript

Published in
IEEE Transactions on Biomedical Engineering

Citation (APA)
Adal, K. M., Van Etten, P. G., Martinez, J. P., Rouwen, K. W., Vermeer, K. A., & van Vliet, L. J. (2017). An Automated System for the Detection and Classification of Retinal Changes Due to Red Lesions in Longitudinal Fundus Images. *IEEE Transactions on Biomedical Engineering*, *PP*(99), 1382-1390. https://doi.org/10.1109/TBME.2017.2752701

An Automated System for the Detection and Classification of Retinal Changes Due to Red Lesions in Longitudinal Fundus Images

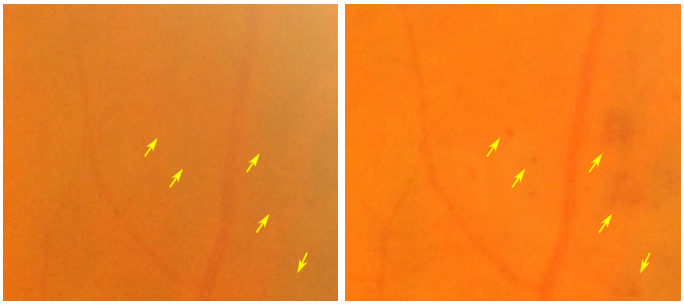
Kedir M. Adal*, Peter G. van Etten, Jose P. Martinez, Kenneth W. Rouwen, Koenraad A. Vermeer, *Member, IEEE* and Lucas J. van Vliet, *Member, IEEE*

***Abstract*—People with diabetes mellitus need annual screening to check for the development of diabetic retinopathy (DR). Tracking small retinal changes due to early diabetic retinopathy lesions in longitudinal fundus image sets is challenging due to intra- and inter-visit variability in illumination and image quality, the required high registration accuracy, and the subtle appearance of retinal lesions compared to other retinal features. This paper presents a robust and flexible approach for automated detection of longitudinal retinal changes due to small red lesions by exploiting normalized fundus images that significantly reduce illumination variations and improve the contrast of small retinal features. To detect spatio-temporal retinal changes, the absolute difference between the extremes of the multiscale blobness responses of fundus images from two time-points is proposed as a simple and effective blobness measure. DR related changes are then identified based on several intensity and shape features by a support vector machine classifier. The proposed approach was evaluated in the context of a regular diabetic retinopathy screening program involving subjects ranging from healthy (no retinal lesion) to moderate (with clinically relevant retinal lesions) DR levels. Evaluation shows that the system is able to detect retinal changes due to small red lesions with a sensitivity of 80% at an average false positive rate of 1 and 2.5 lesions per eye on small and large fields-of-view of the retina, respectively.**

***Index Terms*—Computer-aided detection, fundus images, diabetic retinopathy screening, red lesions, longitudinal analysis.**

I. Introduction

DIABETIC retinopathy is a complication of diabetes mellitus, which progressively damages retinal blood vessels and may result in vision loss and even blindness if not diagnosed and treated adequately. Regular eye examination is necessary for timely detection and treatment of DR at an early stage [1]. The current eye care practice for screening DR involves examination of multiple field fundus images for pathognomonic abnormalities by a trained expert. Depending on the observed retinal abnormalities at the time of the examination, diabetic patients are either scheduled for a follow-up examination or referred to an ophthalmologist for further diagnostic evaluation and possibly treatment. This screening procedure is subjective [2], time consuming and puts a considerable demand on diabetic eye care resources.



(a) Baseline image (b) Follow-up image

Fig. 1: A pair of spatially aligned fundus image patches showing longitudinal retinal change locations (yellow arrows) due to early stage DR lesions between the baseline and follow-up retinal examinations.

Moreover, in addition to examining how far the disease has progressed at the time of examination, the goal of regular DR screening is also to identify patients with a high risk of progression. DR is a progressive disease that results in retinal changes such as the appearance (and the disappearance) of associated lesions such as microaneurysms and hemorrhages (see Fig. 1). Recent studies suggest that in addition to the number of lesions at the time of examination, the dynamics of these lesions is useful to monitor progression of DR [3]–[5]. Automated detection and quantification of retinal changes can thus be an important addition to regular DR screening to objectively assess the disease activity over time for proactively taking appropriate measures. An automated system is also instrumental in patient education, especially in asymptomatic patients. By highlighting and showing DR related retinal changes on a computer display, the patients may have better understanding of their progressing eye condition and the importance of regular checkup and adjustment of their blood sugar level to reduce their risk of developing visual complications.

Automated detection of longitudinal retinal changes from a series of fundus images is challenging for several reasons. Firstly, accurate detection and classification of longitudinal

*K. M. Adal is with Delft University of Technology, Delft, The Netherlands and Rotterdam Ophthalmic Institute, Rotterdam, The Netherlands (correspondence e-mail: K.M.Adal@tudelft.nl).

P. G. van Etten, J. P. Martinez, and K. W. Rouwen are with The Rotterdam Eye Hospital, Rotterdam, The Netherlands

K. A. Vermeer is with Rotterdam Ophthalmic Institute, Rotterdam, The Netherlands.

L. J. van Vliet is with Delft University of Technology, Delft, The Netherlands

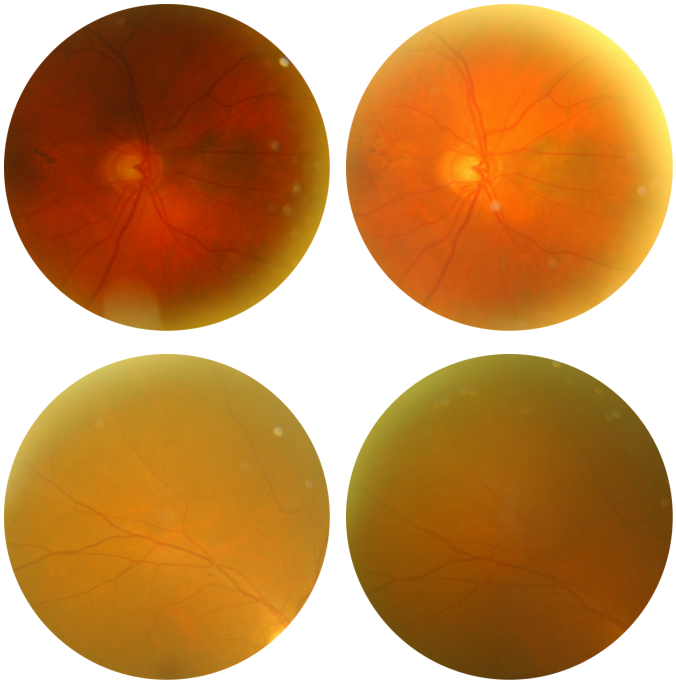


Fig. 2: Examples of fundus image pairs of the same retina (top and bottom rows) captured one year apart during regular DR screening. Significant variation in illumination and acquisition artefacts can be seen in both image pairs.

retinal changes requires a robust algorithm to discriminate clinically relevant changes from changes caused by illumination variation and noise. Fundus imaging is a process that involves careful manual tuning of fundus camera settings, thus color fundus images often suffer from intra- and inter-image variation in luminosity and contrast (see Fig. 2) [6]. Illumination variation coupled with the subtle appearance of early DR lesions renders identifying clinically relevant retinal changes in color fundus images as a very difficult task, even for expert graders.

Secondly, tracking small retinal features, such as microaneurysms and dot hemorrhages, over time requires very high registration accuracy. In order to correctly register fundus images, the nonlinear spatial deformation caused by the projection of the curved surface of the retina onto a flat imaging plane needs to be accounted for [7], [8]. This is challenging due to the sparseness of retinal features that can be used for matching and the limited overlapping region between different fundus fields.

Most of the previous studies on computer-aided detection and diagnosis (CAD) systems for DR screening exclusively aimed at analyzing digital fundus images from a single retinal examination [2], [9]–[22]. A common approach in these studies is to detect early stage retinal abnormalities that are associated with referable DR. Although these CAD systems enable to identify retinal abnormalities for screening DR at the time of examination, they give only limited insight into the activity of the disease since the previous check-up, and thus are not suitable for explicitly monitoring DR.

So far, only a few automated systems have been developed for the detection of longitudinal retinal changes for monitoring DR over time. Examples include systems proposed by Cree et al. [23] and Goatman et al. [24] to detect MAs in longitudinal fluorescein angiogram images. Both systems consist of a method to detect MAs from a region-of-interest centered on the fovea and a registration algorithm to align longitudinal images of the same retina for determining the microaneurysm turnover. Narasimha-Iyer et. al. [25] presented an integrated system for directly detecting and classifying retinal changes from a single-field (macula-centered) color fundus images using a combination of methods for illumination correction, dust removal and segmenting retinal features such as the fovea, optic disc, and blood vessels. The system was later extended to detect changes in vasculature width and appearance/disappearance of lesions [26]. A commercial system also exists for automatically detecting temporal retinal changes due to red-dot-like lesions from a pair of fundus images of the same retina [27], [28].

The main limitation of existing methods for retinal change detection is that they do not address the problems of illumination variation and the space-variant image quality over the entire field-of-view of the retina. Hence, they are not applicable to a large field of the retina, which is required for a comprehensive retinal examination. This is especially a crucial factor when analyzing fundus images in which illumination variation and low image quality affect more than 50% of the FOV as shown in Fig. 2. This paper addresses these problems and presents a robust and flexible multi-stage approach that is applicable to a wide range of fundus fields for automated detection of longitudinal retinal changes due to microaneurysms and dot hemorrhages (small red lesions).

II. Materials and Methods

An overview of the proposed multi-stage approach for automated detection and classification of changes due to small red retinal lesions in longitudinal fundus images is shown in Fig. 3. In the first stage, illumination variation is addressed by normalizing the green channel of each color fundus image for luminosity and contrast variation, thereby improving the visibility of retinal features. Then all the baseline and follow-up sets of normalized four-field fundus images are registered into a common coordinate system using a multi-resolution matching strategy coupled to a hierarchical registration model. In the second stage, spatio-temporal retinal change locations are detected by a novel criterion, blobness measure, based on a multi-scale Laplacian of Gaussian. At the last stage, several local intensity and shape descriptors were extracted from each of the detected change locations and subsequently classified as a change due to a red retinal lesion or no change. Each stage of the proposed approach is described in detail in the following subsections.

A. Dataset

Data for this study was obtained from a regular DR screening program at the Rotterdam Eye Hospital. Four field (macula-centered, optic nerve-centered, superior, and temporal regions) fundus image sets from 81 diabetic eyes that were

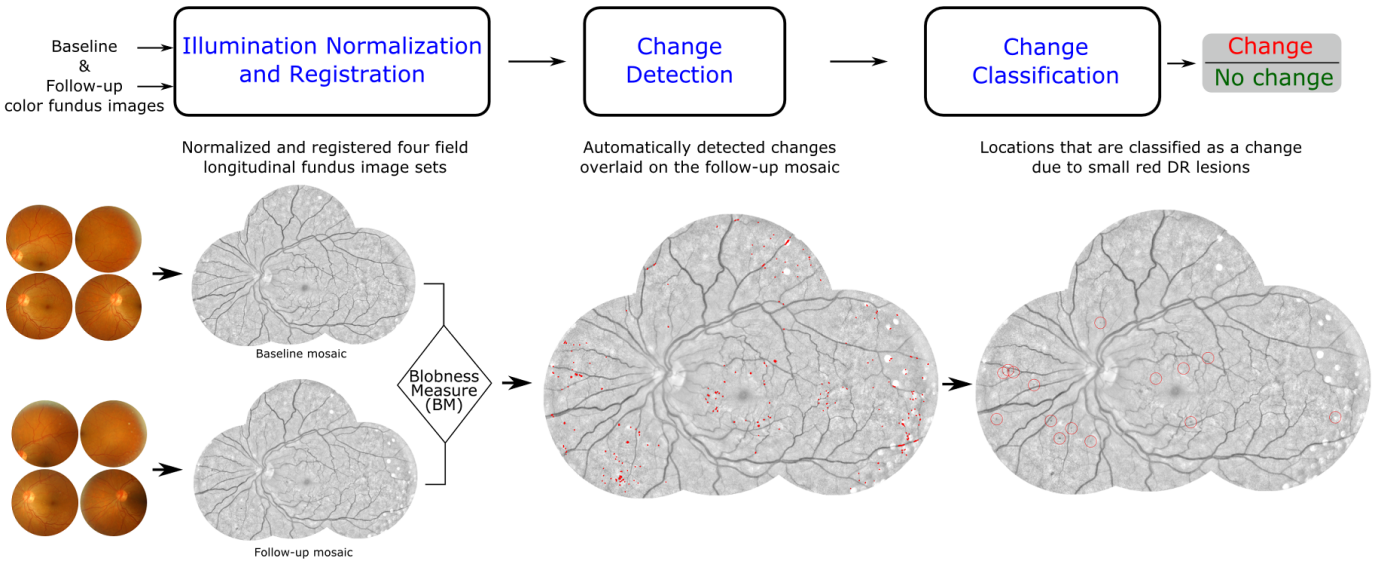


Fig. 3: An overview of the proposed automated system for the detection and classification of longitudinal retinal changes due to red lesions.

acquired for DR screening in 2012 and again in 2013 were used for training (40 eyes) and testing (41 eyes) the proposed approach. All fundus images of 2000×1312 pixels in size were acquired after pupil dilation using a non-mydriatic digital fundus camera (Topcon TRC-NW6S, Tokyo, Japan) with a 45° field of view.

B. Illumination Normalization and Registration

The green channel of digital fundus images is commonly used in automated fundus image analysis because of its higher contrast between retinal features and the background than the red and blue channels. However, the green channel images show considerable variation in luminosity (brightness) and contrast between retinal structures, both within and between images [6]. In a recent work [29], we have addressed this illumination variation by applying an improved version of the luminosity and contrast normalization technique of Foracchia's et al [6]. The normalization was done by using estimates of the local luminosity and contrast from the intensity distribution of the so-called background retina (i.e., the retina excluding features such as vessels, optic disc, and lesions) and subsequently correcting for their variation over the entire retinal image. In order to take into account missing retinal features, especially around the borders of fundus images, the local luminosity and contrast were estimated based on normalized convolution [29]. This results in a normalized retinal image with a uniform illumination pattern within the entire retinal field and improved visibility of fine retinal details (see Fig. 4).

In order to track small retinal features, such as small red lesions, over time, a very high registration accuracy is required. The curved nature of the retinal surface introduces a nonlinear spatial deformation in the process of acquiring fundus images. Therefore, a higher order (quadratic) deformation model is needed for registering fundus images accurately. Over the past decade, several algorithms have been proposed for registration of fundus images [7], [8], [30]–[34]. In this paper, we used a recently introduced robust fundus image registration method that exploits the normalized intensity as well as the structural information of the retinal vasculature [21]. The method aligns retinal vessels based on a multi-resolution matching strategy coupled to a hierarchical registration model with a deformation model of increasing complexity for robust optimization of a global second-order transformation model. The method was successfully applied to register four-field (macula-centered, optic nerve-centered, superior, and temporal) intra- and inter-visit fundus images that capture different parts of the same retinal surface [35].

C. Retinal Change Detection

At early stages, DR is associated with microaneurysms, swellings in small blood vessels that may leak blood into the retina. These lesions commonly appear in color fundus images as small, round dark-red spots (see Fig. 4 and 5).

Due to the resemblance of these lesions to roundish blobs, the Laplacian of Gaussian (LoG) operator is proposed for detecting them in normalized fundus images. The LoG is sensitive to a certain scale and will therefore provide a maximum response at the scale that matches the size of the object to be detected. In order to take into account the reduction of the LoG response with an increase in σ, the scale-normalized LoG operator is defined as [36]

$$\nabla^2_{norm}(\sigma) = \sigma^2 \nabla^2 G(x, y; \sigma), \tag{1}$$

where $\nabla^2 G(x, y; \sigma) = \frac{\partial^2 G(x,y;\sigma)}{\partial x^2} + \frac{\partial^2 G(x,y;\sigma)}{\partial y^2}$ is the LoG operator and $G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$ is a 2D Gaussian function of scale σ.

In longitudinal DR screening, the focus is mainly on identifying regions that have changed due to an appearance or disappearance of retinal lesions between DR checkups. Thus,

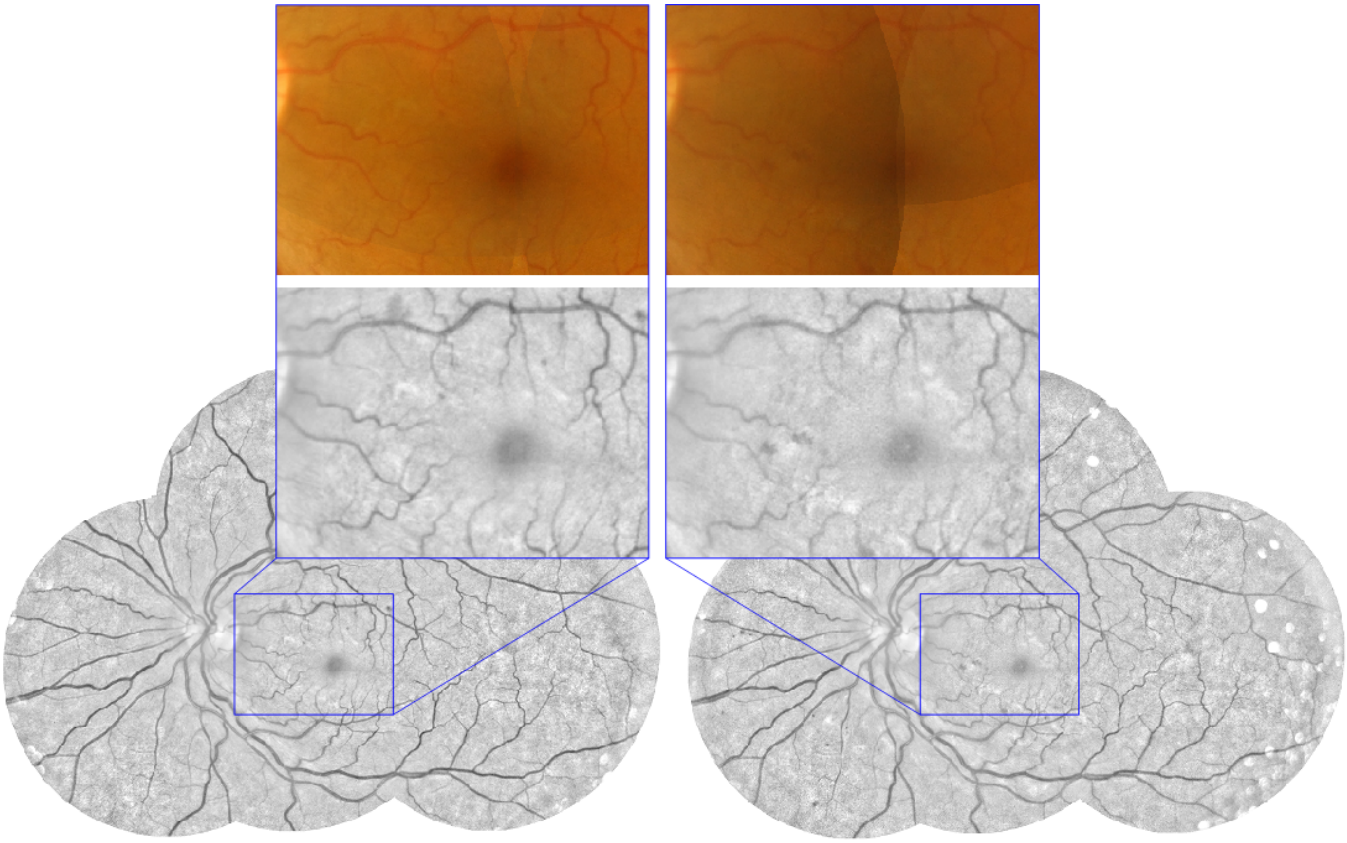


Fig. 4: Examples of baseline (left) and follow-up (right) normalized fundus mosaics produced by registering four-field fundus images captured from a left eye during regular DR screening. The overlaid color and normalized (enlarged) image patches highlight retinal changes due to DR lesions.

given the two time-point fundus images I_{t_1} and I_{t_2}, spatio-temporal changes are detected by first applying the scale-normalized LoG operator at several scales to each of the two time-point images and then comparing the results. To this end, we propose a blobness measure (BM), which is defined as the absolute difference between the extremes of the multiscale blobness responses of fundus images from two time-points.

$$BM(I_{t_1}, I_{t_2}) = |\max_{\sigma} \nabla^2_{norm}(\sigma) * I_{t_1} - \max_{\sigma} \nabla^2_{norm}(\sigma) * I_{t_2}|. \quad (2)$$

In regions that changed due to an appearance or disappearance of retinal lesions between the two time-points, the BM is expected to be significant (see Fig. 6a-6c). A candidate change mask (M) is then obtained by thresholding the BM response at θ_{BM} as

$$M = \{(x,y) | BM(I_{t_1}(x,y), I_{t_2}(x,y)) \geq \theta_{BM}\} \quad (3)$$

In addition, candidate regions that are smaller than the smallest microaneurysm size (3 pixels or $21\mu m$ in diameter in our dataset) were excluded. Figure 6d shows an example change mask extracted from a pair of fundus image patches.

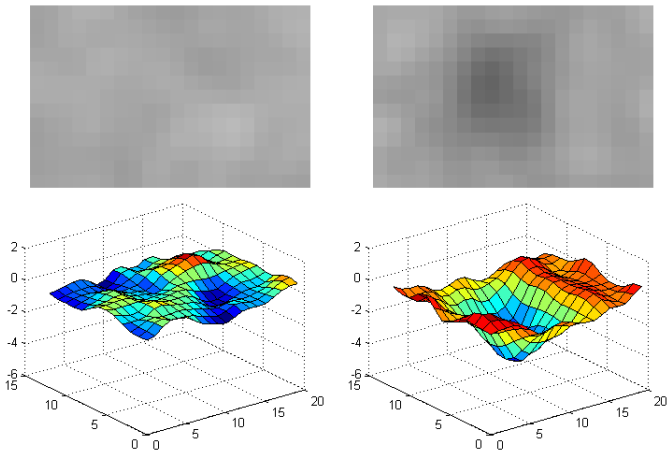
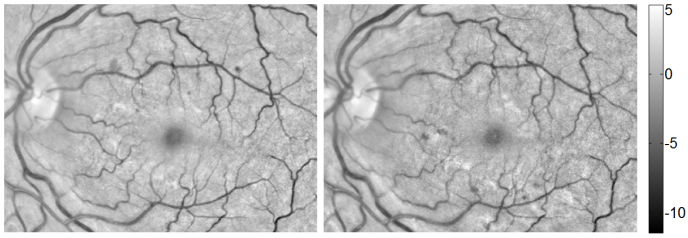


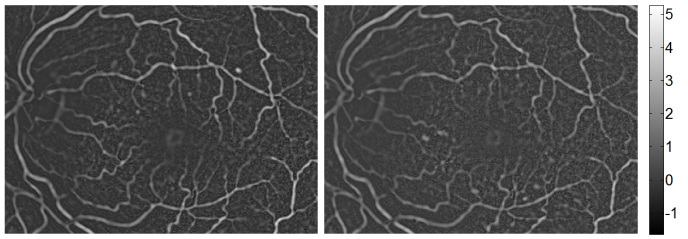
Fig. 5: A pair of spatially-aligned normalized retinal image patches and their surface topographies before (left) and after (right) a small red DR lesion appears.

D. Red Lesion Classification

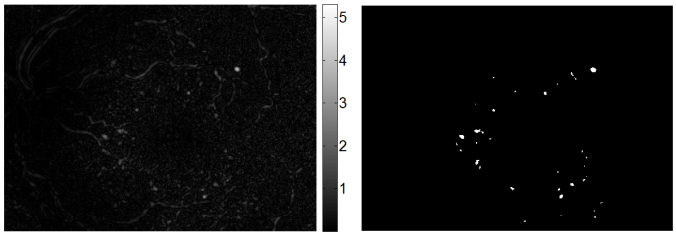
After detecting candidate regions several intensity features, image quality measures, and appearance and shape descriptors were extracted from each candidate region. The complete list of features is



(a) I_{t_1} and I_{t_2}



(b) $\max_\sigma \nabla^2_{norm}(\sigma) * I_{t_1}$ and $\max_\sigma \nabla^2_{norm}(\sigma) * I_{t_2}$



(c) $BM(I_{t_1}, I_{t_2})$ (d) Change mask (M)

Fig. 6: An example of a pair of spatially-aligned normalized image patches (a), the maximum multi-scale blobness responses (b), and the resulting BM values of each pixel (c). Retinal changes due to appearance or disappearance of lesions appear in the BM image as bright white spots. The white spots in (d) indicate the final candidate change mask locations derived from the BM.

1) Intensity features (F_{int}):

- The red and green channel values of the baseline and follow-up images.
- The normalized intensities of the baseline and follow-up images, I_{t_1} and I_{t_2}.
- The difference in luminosity (L) and contrast (C) between the baseline and follow-up images computed as $\triangle L = |L_{t_1} - L_{t_2}|$ and $\triangle C = \sqrt{C^2_{t_1} + C^2_{t_2}}$. L and C are respectively estimated from the instensity distribuition of the retinal background image(excluding vessels, optic disc, and lesions) using the sample mean and standard deviation [29].
- Local power per frequency band of normalized images, which is defined as the extent to which the intensities of retinal regions change rapidly and locally. This is obtained by computing the signal power after band-pass filetring.

A total of 20 intensity features were extracted from each candidate region.

2) Appearance and shape descriptors (F_{asd}):

- The blobness measure (BM) (see eq. 2).
- Histogram of oriented gradients (HOG) [37] computed with a cell size of 5×5 pixels and block size of 4×4. This produced 144 features.
- Scale-adapted speeded up robust features (SURF) [38]. 64 SURF features were extracted.
- The local intensity curvatures (eigenvalues) computed from the second-order Gaussian derivatives computed at $\sigma = 2^{i/2}, i \in \{0, 1, 2, 3, 4, 5\}$.

A total of 221 appearance and shape descriptors features were extracted.

Three classifiers, K-nearest neighbor (KNN), random forests (RF), and a support vector machine (SVM) with a radial basis function (RBF) kernel, were independently used to predict the probability that each candidate region is a change due to a red retinal lesion.

E. Reference Annotation Formation

The reference annotations used for both training and testing the proposed system was gathered from three experts on DR screening (two ophthalmologists and an optometrist). Each of the graders independently annotated the center locations of retinal changes between the baseline and follow-up exam due to small red DR lesions in the fundus mosaics for each eye. The experts were shown both the color and normalized mosaics side-by-side using custom-made software that we developed for lesion annotation. In order to handle inter-grader annotation variability, the reference annotation was defined based on the simultaneous truth and performance level estimation (STAPLE) algorithm [39].

III. Experiments and Results

A. Evaluation metrics

The performance evaluation metrics were the sensitivity (the proportion of correctly detected and classified lesion locations) and average number of false positives per eye. These metrics were computed as

$$Sensitivity = \frac{TP}{TP + FN} \tag{4}$$

$$Average\ FPs\ per\ Eye = \frac{FP}{N}, \tag{5}$$

where TP is the number of true positives, FN is the number of false negatives, FP is the number of false positives and N is the number of eyes in the test set. A detected location is counted as TP if the distance between its centroid and the closest reference annotation is less than 7 pixels. The evaluation metrics were measured for several threshold levels applied to the (prediction) probability assigned to each of the candidate locations by the classifier and the results are summarized using free-response receiver operating characteristics (FROC) curves.

B. Parameter Settings

The settings of the two parameters (σ, θ_{BM}) for the retinal change detection algorithm (eq. 1-3) were optimized based on the training set. The σ values were determined from the relationship between the size (diameter d) of a retinal lesion and the scale at which the lesion response to the scale-normalized LoG operator achieves its maximum. The σ value can be computed as

$$\sigma = \frac{d}{2\sqrt{2}}. \tag{6}$$

The estimated diameter of the retinal lesions in our dataset ranges from 3 to 16 pixels ($21\mu m$ to $112\mu m$) and thus scales of $\sigma = 2^{i/2}, i \in \{0, 1, 2, ..., 5\}$ pixels were applied to eq. 2. The change detection sensitivity and number of false candidates were optimized by varying a range of values $\theta_{BM} \in \{1.0, 1.1, 1.2, 1.3, 1.4, 1.5\}$. A threshold value of $\theta_{BM} = 1.2$ provided the best compromise between sensitivity (which was set to be at least 96%) and average number of false candidates per eye.

The optimal values for the RBF kernel parameters ($C = 2^{0.5}$, $\gamma = 2^{-9}$) of the SVM classifier and for the number of nearest neighbors ($K = 25$) of the KNN classifier were chosen through 10-fold cross-validation on the training set. A grid-search [40] in combination with cross-validation was used to test various (C, γ) pairs. The RF classifier parameters (the number of trees, the number of randomly selected features for splitting) were set based on the out-of-bag (OOB) data error estimate [41]. During each bootstrap the RF classifier sets aside about one-third of the training samples as OOB data and this data is not used for constructing a tree; therefore, it is used internally as a validation set to estimate the classification error.

C. Evaluations

We evaluated the proposed approach for the detection and classification of retinal changes due to red lesions from longitudinal retinal mosaics on both a large and a small field of view of the retina on the test set. The evaluation on a large retinal field was done using the four field fundus mosaics, which consists of the macula, optic-disc, superior, and temporal retinal regions. The number of retinal changes (appearances and disappearances) for each eye in the training and testing set is shown in figure 7. The total number of retinal changes in the training and testing set were 174 and 164, respectively. The evaluation on a small retinal field was done based on the macula-centered fundus images due to their clinical significance. In addition, the performance of each of the two feature types (intensity vs. appearance and shape) paired with each of the three classifiers (KNN, RF, SVM) was evaluated on the large field fundus mosaics.

D. Results

Figure 8 shows the FROC curves for the systems with various classifier and feature combinations. The blue horizontal line indicates the retinal change detection sensitivity (98%) of the proposed approach on the testing set, which acts as an upper bound on the sensitivity of the whole system. The results show that for all tested classifiers the appearance and shape descriptors produced a much better classification result than the intensity features. Moreover, the performance of the classifiers increased when the combined set of features was used. The SVM classifier performed best among the three classifiers and achieved a sensitivity of 80% at an average false positive rate of 2.5 per eye.

The system that performed best on the large field retinal mosaics (SVM with F_{int}+ F_{asd}) was retrained and used to detect retinal changes from a small field of the retina centered on the macula. The performance of the proposed approach is shown in the FROC curve in figure 9. The sensitivity of the retinal change detection algorithm was 97% on macula-centered fundus images. The overall system achieves a sensitivity of 80% at an average false positive rate of 1 per eye.

For each eye in the test set, we visually inspected and analyzed those locations that were detected by our approach but not defined as a DR related change in the reference annotations and thus were counted as false alarms. These locations were found to be in either of the following categories:

- **Dark spots** that resemble tiny red retinal lesions on either

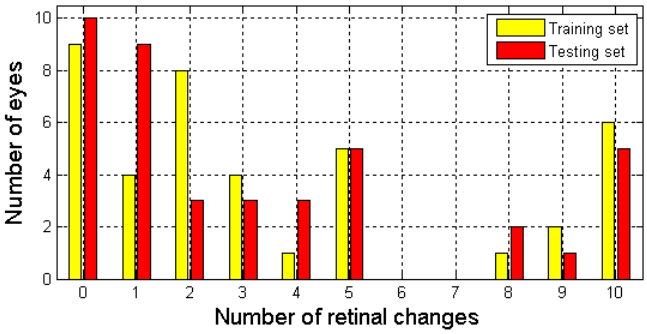


Fig. 7: Distribution of the number of retinal changes in the training and testing sets.

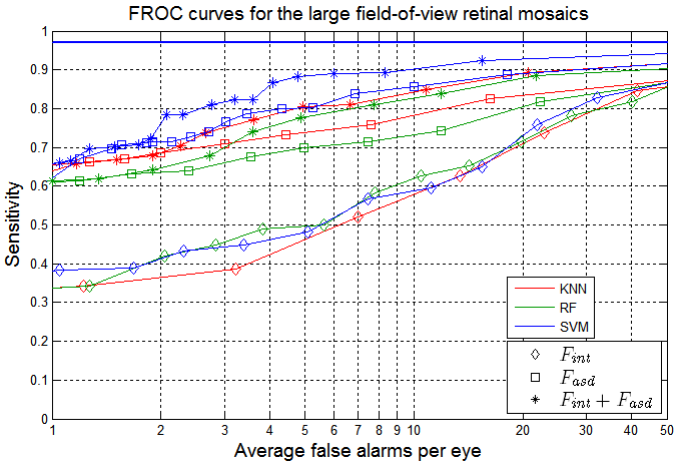


Fig. 8: FROC curves of the proposed systems for the detection and classification of longitudinal retinal changes due to small red lesion applied to large field-of-view retinal mosaics.

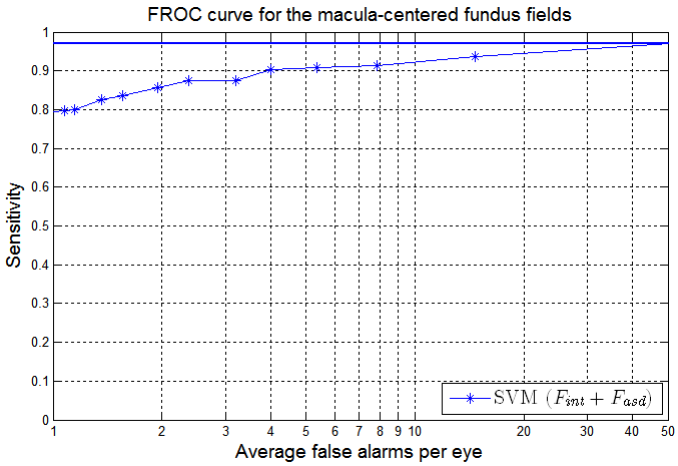


Fig. 9: FROC curve of the proposed system for the detection and classification of longitudinal retinal changes due to small red lesion applied to macula-centered fundus image pairs.

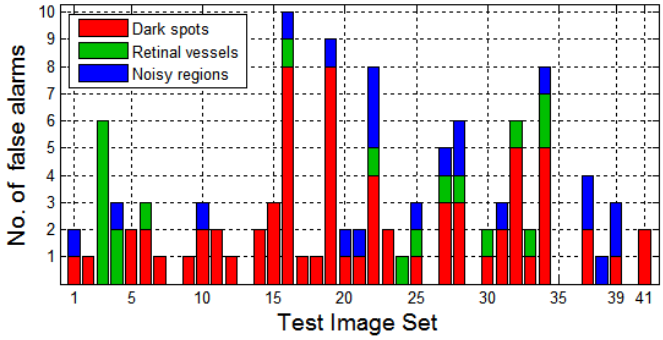


Fig. 10: Histograms depicting the number of false alarms for each category.

the baseline or follow-up mosaics.

- **Retinal vessels** that were affected by illumination artefacts on either the baseline or follow-up mosaics.
- **Noisy regions** that have a low signal-to-noise (SNR) ratio.

Figure 10 shows the number of false alarms in each of the three categories for the 41 test eyes at 0.80 sensitivity. Overall, the results show that dark spots that resemble small red lesions caused 63% of the false alarms. The remaining false alarms were detected either on retinal vessels (18%) or in regions with a low SNR (19%). An example from each of the three categories of false alarms are shown in figure 11.

IV. Discussion and Conclusion

In this paper, we have presented a robust and flexible multi-stage approach for tracking retinal changes due to small red DR lesions such as microaneurysms and dot hemorrhages in longitudinal fundus images. The system was applied to both small and large retinal fields of 81 diabetic eyes. Robustness to intra and inter-image illumination variation was achieved by exploiting fundus images that are normalized for luminosity and contrast over the entire field of view. The improvement in the visibility and contrast of especially small retinal features in the normalized fundus images enabled our approach to track subtle retinal changes, including those that are visually difficult to detect on the color fundus images. A simple and effective criterion for blobness (BM) was defined for detecting spatio-temporal retinal change locations from longitudinal normalized fundus images. The BM can also be easily adapted to other related problems for the detection and tracking of small round objects in a series of registered longitudinal images.

The proposed approach was evaluated in the context of a regular diabetic retinopathy screening program involving subjects ranging from healthy (no retinal lesion) to moderate (with clinically relevant retinal lesions) DR levels. Evaluation was done on both a large field-of-view fundus mosaics, which consisted of the macula, optic nerve, temporal, and superior fields, and a small field-of-view of the retina consisting only of the macula-centered fields. The results show that the system was able to detect retinal changes due to small DR lesions with a sensitivity of 80% from large field fundus mosaics and small field fundus images at an average false positive rate of 2.5 and 1, respectively. In contrast to the small fields, the higher false alarm rate in the large field fundus mosaics is mainly caused by the lower image quality and the presence of significant illumination artefacts, such as white spots (see Fig. 4), on the temporal and superior fundus fields.

Visual inspection of the false alarms suggests that most of them were very similar in appearance and shape to small red DR lesions and thus may well be true positives that were erroneously not included in the reference annotation. Indeed, 49% of the detected dark spot locations were small red lesions that were also annotated by one of the graders. It should also be noted that color fundus images are routinely used by eye care experts in DR screening. Therefore, introducing the normalized images during annotation can help the experts see subtle DR related retinal changes, although their annotations might have been biased towards the color images.

The proposed approach could also be applied to determine the red DR lesion count of individual fundus images provided that reference images with known retinal conditions are available. Automated detection of red DR lesions from single time point images can be very difficult due to the subtle nature of most of the lesions and limited number of lesion pixels. On a publicly available Retinopathy Online Challenge fundus image dataset [42], the top ranking method for automated detection of red lesions from individual images achieved a sensitivity of 53% at an average of 2 false alarms per image [16]. By incorporating reference images and analyzing spatio-temporal change locations, our approach could be applied to detect and determine red lesion count with a higher sensitivity.

Automated detection and quantification of longitudinal retinal changes can be an important addition to regular DR screening. The detected retinal changes can be used for making objective and quantitative analysis of DR progression as well as for more efficient human grading and patient education by highlighting DR related changes since the previous visit.

Acknowledgment

This work is financially supported by CZ Fonds, Stichting Blindenhulp and Stichting Wetenschappelijk Onderzoek het

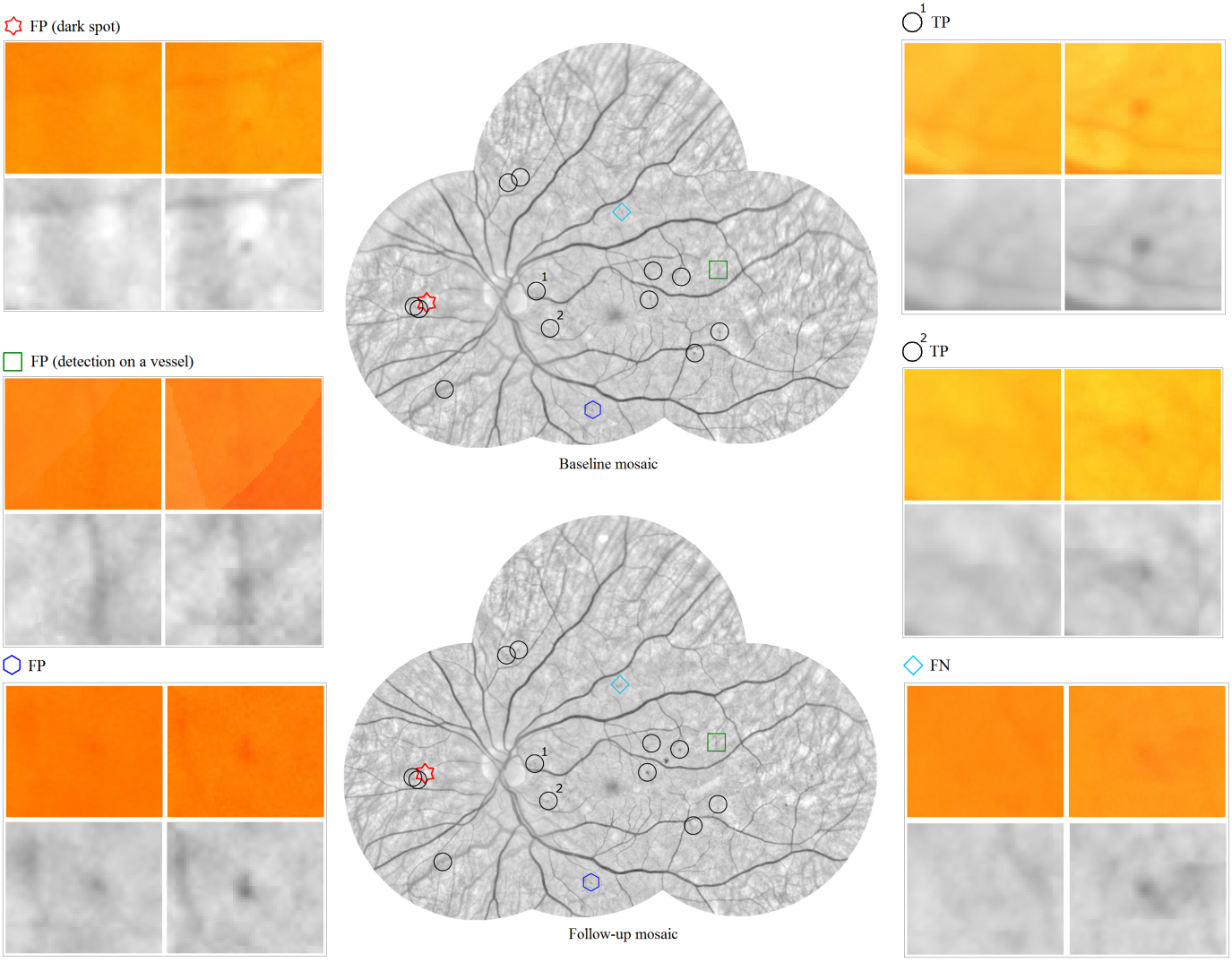


Fig. 11: An example of the output of the proposed change detection and classification approach at a sensitivity of 80%. The zoomed in color and normalized image patches for the baseline (left) and follow-up (right) indicated by different shapes represent the true positives (TPs), false positives (FPs) and false negative (FN) locations.

Oogziekenhuis.

References

[1] Early Treatment Diabetic Retinopathy Study Research Group, "Early photocoagulation for diabetic retinopathy: ETDRS report number 9," *Ophthalmology*, vol. 98, no. 5, pp. 766–785, 1991.

[2] M. D. Abràmoff *et al.*, "Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes," *Diabetes care*, vol. 31, no. 2, pp. 193–198, 2008.

[3] S. Nunes *et al.*, "Microaneurysm turnover is a biomarker for diabetic retinopathy progression to clinically significant macular edema: findings for type 2 diabetics with nonproliferative retinopathy," *Ophthalmologica*, vol. 223, no. 5, pp. 292–297, 2009.

[4] C. Haritoglou *et al.*, "Microaneurysm formation rate as a predictive marker for progression to clinically significant macular edema in non-proliferative diabetic retinopathy," *Retina*, vol. 34, no. 1, pp. 157–164, 2014.

[5] J. Cunha-Vaz *et al.*, "Phenotypes and biomarkers of diabetic retinopathy," *Progress in retinal and eye research*, vol. 41, pp. 90–111, 2014.

[6] M. Foracchia *et al.*, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis*, vol. 9, no. 3, pp. 179–190, 2005.

[7] A. A. Mahurkar *et al.*, "Constructing retinal fundus photomontages. a new computer-based method." *Investigative Ophthalmology & Visual Science*, vol. 37, no. 8, pp. 1675–1683, 1996.

[8] A. Can *et al.*, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 3, pp. 347–364, 2002.

[9] M. Niemeijer *et al.*, "Automatic detection of red lesions in digital color fundus photographs," *IEEE Transactions on medical imaging*, vol. 24, no. 5, pp. 584–592, 2005.

[10] T. Walter *et al.*, "Automatic detection of microaneurysms in color fundus images," *Medical image analysis*, vol. 11, no. 6, pp. 555–566, 2007.

[11] M. Niemeijer *et al.*, "Information fusion for diabetic retinopathy cad in digital color fundus photographs," *IEEE Transactions on Medical Imaging*, vol. 28, no. 5, pp. 775–785, 2009.

[12] A. D. Fleming *et al.*, "The role of haemorrhage and exudate detection in automated grading of diabetic retinopathy," *British Journal of Ophthalmology*, vol. 94, no. 6, pp. 706–711, 2010.

[13] M. Abràmoff *et al.*, "Retinal imaging and image analysis," *Biomedical Engineering, IEEE Reviews in*, vol. 3, pp. 169–208, 2010.

[14] M. Niemeijer *et al.*, "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 185–195, 2010.

[15] G. Quellec *et al.*, "Optimal filter framework for automated, instantaneous detection of lesions in retinal images," *IEEE Transactions on medical imaging*, vol. 30, no. 2, pp. 523–533, 2011.

[16] B. Antal and A. Hajdu, "An ensemble-based system for microaneurysm detection and diabetic retinopathy grading," *IEEE transactions on biomedical engineering*, vol. 59, no. 6, pp. 1720–1726, 2012.

[17] A. Rocha *et al.*, "Points of interest and visual dictionaries for automatic retinal lesion detection," *IEEE transactions on biomedical engineering*, vol. 59, no. 8, pp. 2244–2253, 2012.

[18] L. Giancardo *et al.*, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," *Medical image analysis*, vol. 16, no. 1, pp. 216–226, 2012.

[19] L. Tang *et al.*, "Splat feature classification with application to retinal hemorrhage detection in fundus images," *IEEE Transactions on Medical Imaging*, vol. 32, no. 2, pp. 364–375, 2013.

[20] M. D. Abràmoff *et al.*, "Automated analysis of retinal images for detection of referable diabetic retinopathy," *JAMA ophthalmology*, vol. 131, no. 3, pp. 351–357, 2013.

[21] K. M. Adal *et al.*, "Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning," *Computer methods and programs in biomedicine*, vol. 114, no. 1, pp. 1–10, 2014.

[22] I. N. Figueiredo *et al.*, "Automated lesion detectors in retinal fundus images," *Computers in biology and medicine*, vol. 66, pp. 47–65, 2015.

[23] M. Cree *et al.*, "A fully automated comparative microaneurysm digital detection system." *Eye (London, England)*, vol. 11, pp. 622–628, 1996.

[24] K. A. Goatman *et al.*, "Automated measurement of microaneurysm turnover," *Investigative ophthalmology & visual science*, vol. 44, no. 12, pp. 5335–5341, 2003.

[25] H. Narasimha-Iyer *et al.*, "Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy," *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 6, pp. 1084–1098, 2006.

[26] ——, "Integrated analysis of vascular and nonvascular changes from color retinal fundus image sequences," *IEEE transactions on biomedical engineering*, vol. 54, no. 8, pp. 1436–1445, 2007.

[27] C. M. Oliveira *et al.*, "Improved automated screening of diabetic retinopathy," *Ophthalmologica*, vol. 226, no. 4, pp. 191–197, 2011.

[28] J. Cunha-Vaz *et al.*, "Computer-aided detection of diabetic retinopathy progression," in *Digital Teleretinal Screening*. Springer, 2012, pp. 59–66.

[29] K. M. Adal *et al.*, "A hierarchical coarse-to-fine approach for fundus image registration," in *Biomedical Image Registration*. Springer, 2014, pp. 93–102.

[30] N. Ritter *et al.*, "Registration of stereo and temporal images of the retina," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 5, pp. 404–418, 1999.

[31] G. K. Matsopoulos *et al.*, "Automatic retinal image registration scheme using global optimization techniques," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 3, no. 1, pp. 47–60, 1999.

[32] C. V. Stewart *et al.*, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 11, pp. 1379–1394, 2003.

[33] T. Chanwimaluang *et al.*, "Hybrid retinal image registration," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 10, no. 1, pp. 129–142, 2006.

[34] I. N. Figueiredo *et al.*, "Automated retina identification based on multiscale elastic registration," *Computers in biology and medicine*, vol. 79, pp. 130–143, 2016.

[35] K. M. Adal *et al.*, "Accuracy assessment of intra-and intervisit fundus image registration for diabetic retinopathy screening," *Investigative ophthalmology & visual science*, vol. 56, no. 3, pp. 1805–1812, 2015.

[36] T. Lindeberg, "Feature detection with automatic scale selection," *International journal of computer vision*, vol. 30, no. 2, pp. 79–116, 1998.

[37] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.

[38] H. Bay *et al.*, "Speeded-up robust features (surf)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[39] S. K. Warfield *et al.*, "Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation," *IEEE transactions on medical imaging*, vol. 23, no. 7, pp. 903–921, 2004.

[40] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[41] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[42] M. Niemeijer *et al.*, "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 185–195, 2010.