Fully Automatic Segmentation of Fluorescein Leakage in Subjects With Diabetic Macular Edema

Hossein Rabbani,[1,2] Michael J. Allingham,[1] Priyatham S. Mettu,[1] Scott W. Cousins,[1] and Sina Farsiu[1,3]

[1]Department of Ophthalmology, Duke University Medical Center, Durham, North Carolina, United States
[2]Medical Image and Signal Processing Research Center, Isfahan University of Medical Sciences, Isfahan, Iran
[3]Department of Biomedical Engineering, Duke University, Durham, North Carolina, United States

PURPOSE. To create and validate software to automatically segment leakage area in real-world clinical fluorescein angiography (FA) images of subjects with diabetic macular edema (DME).

METHODS. Fluorescein angiography images obtained from 24 eyes of 24 subjects with DME were retrospectively analyzed. Both video and still-frame images were obtained using a Heidelberg Spectralis 6-mode HRA/OCT unit. We aligned early and late FA frames in the video by a two-step nonrigid registration method. To remove background artifacts, we subtracted early and late FA frames. Finally, after postprocessing steps, including detection and inpainting of the vessels, a robust active contour method was utilized to obtain leakage area in a 1500-μm-radius circular region centered at the fovea. Images were captured at different fields of view (FOVs) and were often contaminated with outliers, as is the case in real-world clinical imaging. Our algorithm was applied to these images with no manual input. Separately, all images were manually segmented by two retina specialists. The sensitivity, specificity, and accuracy of manual interobserver, manual intraobserver, and automatic methods were calculated.

RESULTS. The mean accuracy was 0.86 ± 0.08 for automatic versus manual, 0.83 ± 0.16 for manual interobserver, and 0.90 ± 0.08 for manual intraobserver segmentation methods.

CONCLUSIONS. Our fully automated algorithm can reproducibly and accurately quantify the area of leakage of clinical-grade FA video and is congruent with expert manual segmentation. The performance was reliable for different DME subtypes. This approach has the potential to reduce time and labor costs and may yield objective and reproducible quantitative measurements of DME imaging biomarkers.

Keywords: leakage segmentation, diabetic macular edema, fundus fluorescein angiography, nonrigid registration

Diabetic retinopathy is the leading cause of vision loss in working-age adults, affecting a large subset of the over 24 million diabetics in the United States[1,2] and an even greater number worldwide. Diabetic macular edema (DME) affects over 25% of diabetics with 20 years or more duration,[3] and is the primary cause of central vision loss due to diabetic retinopathy. Diabetic macular edema results from a combination of pathologic leakage from damaged retinal microvasculature and insufficient clearance of plasma by Müller and retinal pigment epithelial cells.[4,5] Vascular leakage and intraretinal fluid accumulation are imaged clinically using fundus fluorescein angiography (FA).

While noninvasive optical imaging systems such as optical coherence tomography (OCT) provide valuable morphologic information and are useful to monitor DME and its response to treatment,[6] FA remains essential for diagnosis and characterization of DME disease. Fluorescein angiography offers critical biological information such as location, intensity, and leakage source; and leakage area as measured by FA continues to be a relevant secondary endpoint in major studies of DME treatment.[7] In addition, various subtypes of DME have been proposed based on differences in the pattern of fluorescein leakage as seen by FA.[8] For example, focal leakage manifests as discrete foci of leakage on early FA frames and corresponds to microaneurysms (MAs). In contrast, the diffuse subtype is characterized by generalized leakage prominent on late FA frames without a discretely identifiable source. Eyes with DME can demonstrate either leakage pattern, or more commonly, a mixture of both.[9]

Identification of DME subtypes by FA has potential to guide therapy and monitor disease activity. While reproducible quantitative and qualitative analysis of FA is possible by experienced graders in the setting of a formal imaging reading center, its use for subtyping in the clinical setting is hindered by the subjective nature of FA interpretation. Accordingly, there has been longstanding interest in objective methods for quantification of leakage by FA. While several investigators have utilized automated segmentation for automatic analysis of FA,[10–21] MA detection,[22–27] extraction of vessels,[15,18,19,28] and foveal avascular zone (FAZ) detection,[14,29–33] relatively few algorithms have been focused on automated leakage detection or quantification.[10,34–39]

TABLE 1. FOV of FA Images and Video in This Study

Data	FOV of the Early-Phase FA Videos	FOV of the Late-Phase Images in ART Mode
Diffuse 1	30	55
Diffuse 2	55	55
Diffuse 3	55	55
Diffuse 4	55	55
Diffuse 5	55	55
Diffuse 6	55	55
Diffuse 7	30	30
Focal 1	55	55
Focal 2	35	35
Focal 3	55	55
Focal 4	30	55
Focal 5	55	55
Focal 6	30	30
Focal 7	30	35
Focal 8	55	55
Focal 9	55	55
Focal 10	55	55
Mixed 1	55	55
Mixed 2	30	30
Mixed 3	55	55
Mixed 4	30	35
Mixed 5	30	30
Mixed 6	30	30
Mixed 7	30	55

Martinez-Costa et al.[35] have published a method for detection of macular angiographic leakage due to retinal vein occlusion. The foveal center is manually detected, and then images are aligned automatically. Pixels with a statistically high increment in gray level along the sequence within the closest area to the fovea center are segmented as leakage. Another method by Cree et al.[36] assumes that captured images are composed of two functions, one describing the true underlying image and the other the incurred degradation due to uneven illumination or occluded optical pathways. Any leakage of fluorescein dye is then detected by analyzing the restored data and finding areas of the image that do not have normal fluorescence intensity attenuation. The exponential model of fluorescein decay utilized by Cree et al.[36] is an extension of the linear model used by Philips et al.[37,38] In contrast, other researchers claim that the intensity profile of the hyperfluorescent region is not entirely predictable,[39] especially in cases of late filling vasculature, scars caused by laser surgery, or late staining of the optic nerve head. The obtained temporal profiles in the work of Berger,[40] after using a polynomial warping algorithm for FA registration, also show that simple models are not able to correctly match the intensity profile of the hyperfluorescent regions.

To address this problem, Buchanan and Trucco[39] utilized (1) contextual knowledge and (2) spatiotemporal features exploiting the evolution of intensity levels over the sequences of ultra-widefield retinal angiograms to train an AdaBoost algorithm. More recently, El-Shahawy et al.[41] modeled manually cropped macular image in the early frames by a two-dimensional Gaussian surface, which is then subtracted from the corresponding area in late frames to segment the leakage area using a Gaussian mixture model classification algorithm. This algorithm analyzes only one early frame and one late frame, and along with the previously noted studies, uses rigid phase correction registration. All these noted methods either use rigid registration[35–39,41] (the shortcomings of which will be experimentally proven for our problem) or require manual inputs[35,37,38,41] (e.g., in the registration step or for fovea detection).

In this paper, we present a fully automated image segmentation algorithm (which does not require manual inputs) for reproducible and accurate quantification of leakage area in DME. An exciting characteristic of our algorithm is its applicability to real-world clinical images, which often include low-quality images with various sources of outliers, without requiring any manual input.

METHODS

Study Subjects

This study was approved by the Duke University Health System Institutional Review Board (IRB) in accordance with Health Insurance Portability and Accountability Act (HIPAA) regulations and the standards of the 1964 Declaration of Helsinki. Twenty-four eyes of 24 subjects were included in the study. Only images obtained from the transited eye were analyzed. In order to be included, subjects had to be diagnosed with DME based on clinical exam, FA, and OCT imaging. Exclusion criteria included other causes of macular edema, globally poor image quality (due to media opacity or patient cooperation), missing early- or late-frame images, or photographer error that made accurate segmentation even by manual graders impossible in the opinion of the expert graders. In order to test the performance of the algorithm over a wide spectrum of DME subtypes, efforts were made to include representative subjects with predominantly focal, predominantly diffuse, and mixed pattern leakage (as determined by expert clinicians) in the study.

Data Acquisition

Expert clinicians retrospectively identified FA images obtained during routine clinical care at the Duke Eye Center. All images were obtained using a Heidelberg Spectralis 6-mode HRA/OCT unit (Heidelberg Engineering, Heidelberg, Germany). The first minute of the study was captured in movie mode using the high-resolution setting (4.7 frames per second), and subsequent late-phase images were captured as single images in ART mode (averaging nine images). Each grayscale image in the sequence was composed of 768- × 768-pixel images. The FOVs of the early movie and the late-phase images were 30°, 35°, and 55° (Table 1). Following acquisition, image files were deidentified and exported in E2E format for further analysis.

Image Processing Algorithms

A block diagram of our proposed method for leakage detection from FA images of DME patients is shown in Figure 1. The first step in our algorithm is accurate registration of the FA image sequence for each patient, where we register a set number of frames in the video (called registered frames) to one reference frame in the sequence. After accurate registration of the FA sequence, we estimate the normalized difference between the early and late FA images. After several postprocessing steps including detection and inpainting of vessel regions, we find an initial estimate of the leakage area. Finally, we utilize the robust active contour method[42] to accurately detect the boundaries of the leakage region. These steps are discussed in more detail in the following subsections.

Registration. Accurate registration is a critical step because (1) fluorescence level in FA images is different from one subject to another and (2) nonleakage areas (e.g., vessels) also fluoresce. Since the contrast agent accumulates slowly, the leakage area appears most prominently in the later frames of the FA video sequence, as opposed to MAs, vessels, or laser scars, which are more prominent in earlier frames. Thus, a logical approach for detecting the actual leakage area is to
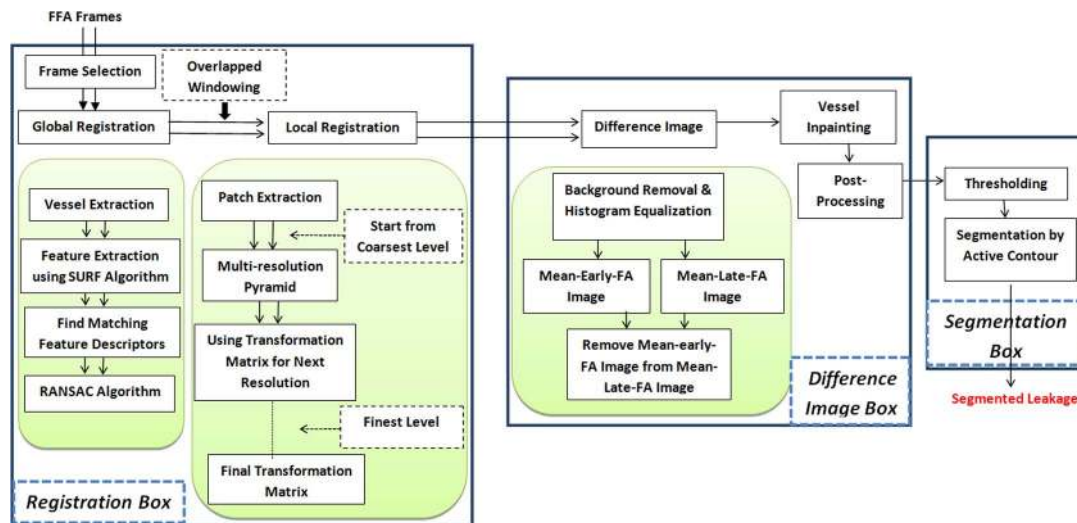
FIGURE 1. Block diagram of proposed method for segmentation of fluorescein leakage areas from FA images of DME patients. In *Registration Box*, selected frames are registered together using a two-step registration method including global and local registration. Two normalized mean early and late frames produced after registration are subtracted in the next stage (*Difference Image Box*). Finally, after thresholding and applying the Chan-Vese segmentation method, segmented leakage is extracted (*Segmentation Box*).

compare the fluorescence levels of geographically similar areas of the retina at different time points.

Registration of an FA sequence, which may span a few minutes, is in general a challenging problem since (1) global and local illumination of frames in an FA video (spanning up to a few minutes) cannot be considered constant; (2) MAs, leakage, and vessels appear and disappear throughout the video; and (3) interframe motion cannot be modeled as rigid (see discussion of Multiresolution Nonrigid Local Registration and Fig. 5 below).

This problem is even more challenging for datasets from a real-world clinical setting (as opposed to a controlled experiment) due to the following issues: (1) different FOVs in FA videos in the same clinical practice (e.g., 30°, 35°, and 55° FOVs); (2) severe distortion of images due to eye movement and blinking; and (3) obstructed view or high levels of noise in a selected number of frames (Fig. 2).

To address these problems, several algorithms with varied levels of success have been proposed through the years.[21,43–51] In our method, to accurately register relatively low-quality clinical FA images, we utilize a two-step nonrigid registration approach: a robust global vessel-based registration method based on the RANdom SAmpling & Consequence (RANSAC) algorithm,[52] followed by a more accurate nonrigid intensity multiresolution registration of FA images.

Frame Selection. The first step of our registration algorithm is removing corrupted frames (especially due to eyelid twitching, blinking, and exceptionally high noise levels) from the registration process (Fig. 2). We achieve this by removing frames with a correlation less than 0.7 with the last frame from the registration process (Fig. 3).

Global Rigid Registration. Once the FA sequence is pruned of the outlier frames, we find a pilot global transform that registers the remaining frames. Our global registration algorithm is based on finding a geometric transformation corresponding to the matching point pairs using a variant of the RANSAC method called the statistically robust M-estimator SAmple Consensus (MSAC) algorithm.[53] The iterative RANSAC method estimates parameters of a mathematical model from a set of observed data that are contaminated with outliers. In MSAC the cost function is modified, whereas inliers are scored according to their fitness to the model while the outliers are given a constant weight. In

order to find the matching point pairs, we first roughly segment the vessels in each image. While virtually any vessel detection algorithm can be employed for this task,[28,54–56] in this paper we use the exploratory Dijkstra forest algorithm of Estrada et al.[54] In this method, after preprocessing, in each iteration the best unvisited vessel pixel in the image is chosen as a starting point for a dynamic-programming exploration of the unvisited part of the image, which results in a new tree in the growing forest of vessels. A threshold is chosen as stop criterion, which stops forest growth when the best unvisited vessel pixel is worse than this threshold.

After this pilot vessel detection step, we utilize the scale and rotation-invariant interest point detector/descriptor Speeded-Up Robust Features (SURF) on the binary vessel map to extract blob features.[57–60] A blob is a region with a (relatively) constant value in properties such as brightness or color compared to areas surrounding that region, which can be utilized as a salient point for registration. In SURF, the determinant of Hessian (DoH) is utilized as the blob detector, computed from the sum of the Haar wavelet response around the point of interest. Figure 4e shows the output of blob detection for two FA frames of a DME patient, and the strongest SURF features are shown in Figure 4f. Next, outliers in blob maps are excluded by using the MSAC algorithm.[58–60] Finally, the remaining blob regions are matched by finding a geometric transformation based on an affine model. This transform was estimated using the estimateGeometricTransform function in MATLAB (MathWorks, Natick, MA, USA) with the parameters of maximum distance threshold, maximum number of random trials for finding the inliers, and desired confidence (in percentage) for finding the maximum number of inliers set at 5, 1000, and 99, respectively. Figure 4 shows an example of global rigid registration between two FA frames of a DME patient. Although global registration improves spatial matching of similar regions in an FA sequence, Figures 5a and 5c show that in some regions further refinement steps are necessary.

Multiresolution Nonrigid Local Registration. To improve the gross global registration results of the previous subsection, we utilize patch-based local registration. After the pilot global registration step, we focus on analyzing local 40- × 40-pixel rectangular patches centered at similarly indexed pixels in the
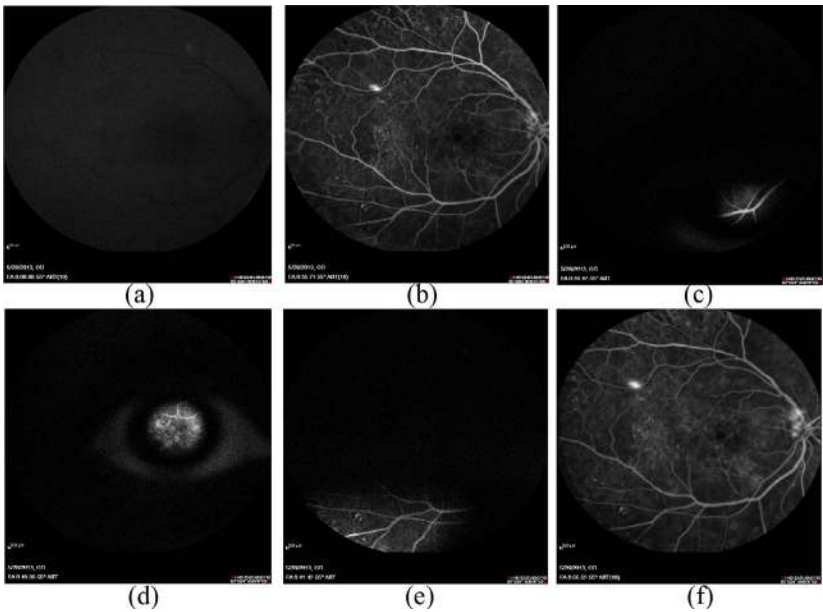
FIGURE 2. Example individual frames of an FA video in our dataset demonstrating the variability of image quality and frequent outliers of FA images captured in a real-world clinical setting. Outlier frames can appear at any time point, complicating development of fully automated software for leakage quantification. (**a**) A low-intensity frame at time point $8''$. (**b**) A frame with acceptable quality at time point $35''$. (**c–e**) Completely unusable (outlier) frames at time points $39''$, $40''$, $41''$. (**f**) A frame with acceptable quality at time point $56''$. The correlations of these six frames to the last frame are 0.61, 0.84, 0.43, 0.44, 0.46, 0.99.

reference and registered images. We use the intensity multi-resolution registration method (implemented utilizing MAT-LAB's imregister function) on the corresponding local patches. To achieve optimal results, in each patch we utilized a multiresolution decomposition approach, with three resolu-tion scales, and iterated 100 times in each pyramidal scale. This procedure can be repeated to obtain the registration param-eters of all pixels. However, we empirically found out that for faster registration, we needed to register only one out of every 20 pixels and used the nearest neighbor coefficients for the
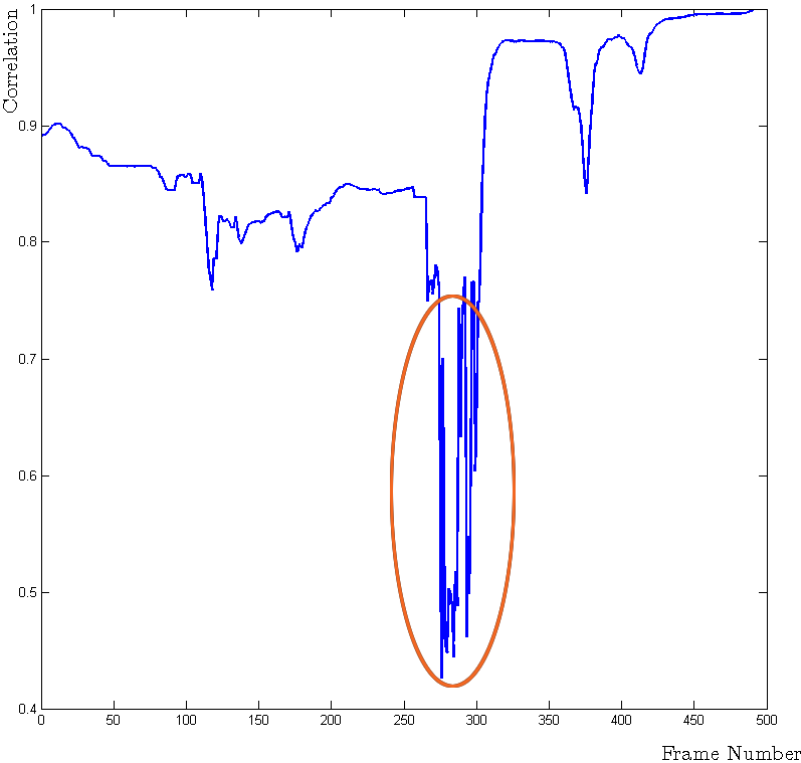

FIGURE 3. Correlation of the 500 frames in the FA sequence (start point is second 11 and end point is second 65) of Figure 2 with the last frame of that sequence. Corrupted frames (corresponding to *orange circle*) with low-correlation values are treated as outliers and are excluded from analysis.
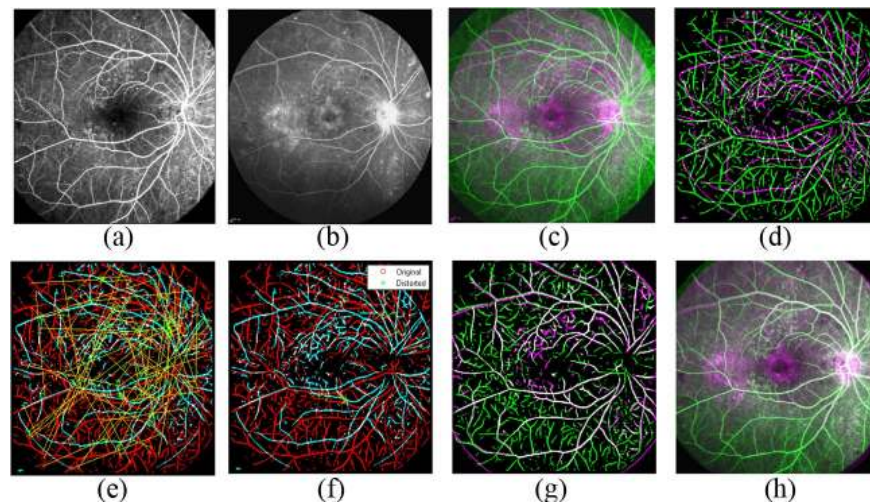
FIGURE 4. An illustrative example of the global (rigid) registration steps for averaged early and late frames of a DME patient. (a) Mean early FA frame. (b) Late FA frame. (c) Unregistered images overlaid. (d) Unregistered vessels overlaid. (e) Initial SURF features of the two frames overlaid. (f) Strongest SURF features overlaid. (g) Rigidly registered vessels. (h) Rigidly registered images. Perfectly registered vessels appear in *white* in (g) and (h).

rest of the pixels. Figures 5b and 5d show the effectiveness of the proposed technique in correcting the slight misalignments in the rigid registration step.

Background Normalization. Following injection of the fluorescent dye, vessels appear in the earlier frames of the FA sequence, followed by MAs, and then leakage areas. In later FA frames, leakage areas are amplified while vessel and MA luminance are attenuated (i.e., early frames show vessels; middle frames show vessels and MAs; and late frames show vessels, MAs, and leakages). Thus, by comparing the FA images captured at different time points, leakage areas can be distinguished from other bright areas in the image. We implement such a background normalization process in the following three steps.

Pilot Background Normalization. Imaging conditions often vary during acquisition of a single FA sequence, which can take several minutes. For example, the incident angle of the laser beam may be different at different time points. Alternately, features such as vessels are attenuated in the later frames as compared to the frames appearing in the middle of the sequence. Thus, the background intensity of the image at local and global scales might be different for different images in a sequence, requiring intensity normalization across all frames. An initial step for intensity normalization is to estimate and subtract the background of each frame. We achieve this by subtracting a morphologically opened variant of each image from itself. Opening in grayscale images is defined as the erosion of image $f(x, y)$ by the structuring element[61] b followed by the dilation of
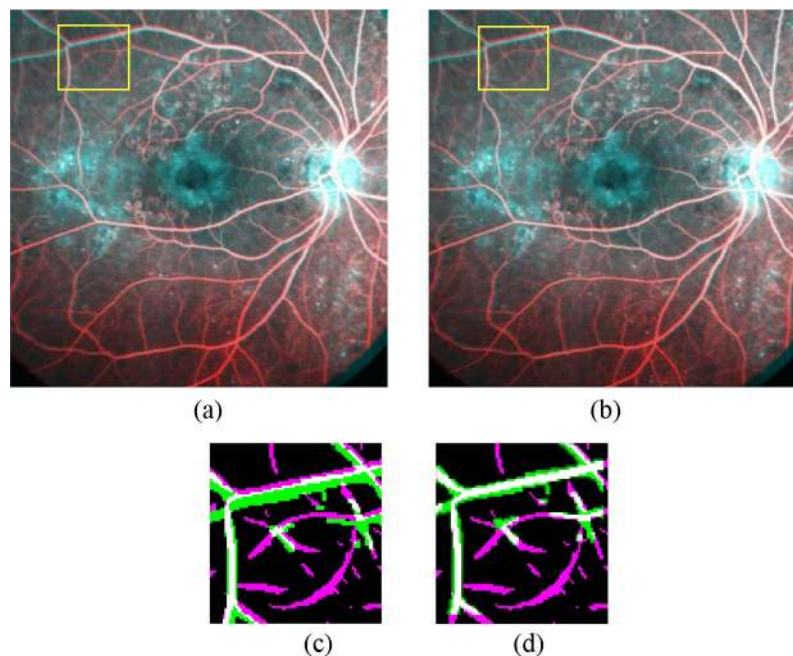

FIGURE 5. Comparison between the results of global rigid registration and nonrigid registration for the image in Figure 4. (a) Overlay of the rigidly registered images. (b) Overlay of the nonrigidly registered images. (c, d) Segmented vessels in the *yellow square section* of (a, b), respectively, where *white* indicates better matching.
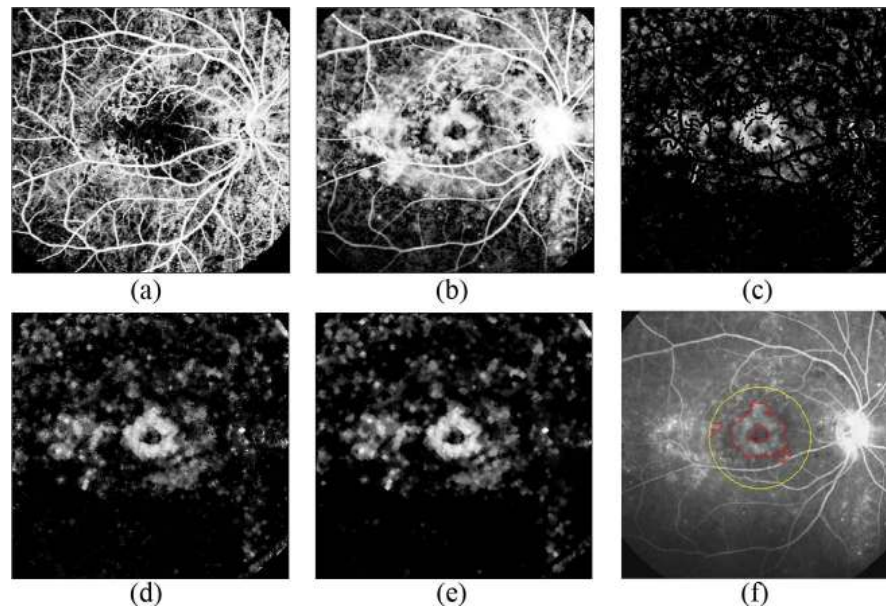
FIGURE 6. Background normalization steps for the image in Figure 4. (a) Pilot background normalized mean early FA frame. (b) Pilot background normalized late FA frame. (c) Pilot vessel and MA removed frame attained by subtracting (b) from (a). (d) Vessel inpainted frame. (e) Removing small objects. (f) Automatically segmented leakage in the 1500-μm-radius ROI.

the result with b. In our implementation, the erosion and dilation operators are defined as $min_{(s,t)\in b}\{f(x + s,\ y + t)\}$ and $max_{(s,t)\in b}\{f(x - s,\ y - t)\}$, respectively, where b is a flat, disk-shaped structuring element with a radius of 20 pixels. Such a relatively large structuring element decreases the intensity of bright features (e.g., vessels and leakage) in our FA images while having a relatively negligible effect on dark features (e.g., FAZ). Thus, by subtracting the opened version of an image from itself, we improve the background intensity uniformity across all images in a sequence. To further improve background uniformity, after background removal we adjust the gray level of each image by local histogram equalization.[62] As an example, Figure 6a is the background-normalized version of Figure 4a.

Pilot Vessel and MA Removal. We accentuate the leakage area in the late FA images by subtracting other fluorescing features, which appear more prominently in earlier frames, such as vessels and MAs. However, individual early FA images are often dominated by image acquisition noise. Thus, instead of subtracting individual frames, we use two representative frames: the averaged early and late frames. The averaged early frame is created by averaging frames 70 to 140. By subtracting mean early FA from late FA image, vessels and MAs in most regions will be significantly attenuated, while leakage areas will be less affected (Fig. 6, first row).

Vessel Masking and Postprocessing. While the previous step eliminates larger vessels, it occasionally fails to remove smaller ones. Moreover, removing vessels located inside a leakage region partitions a continuous leakage area into critically smaller (and undetectable) regions (Fig. 6c). We address this problem by creating an auxiliary image in a two-step set of morphologic operations:

- Removing small objects (e.g., small vessel branches) by applying an opening operation utilizing a disk-shaped structuring element with a radius of 2 pixels; and
- Inpainting the removed vessels by dilating, followed by eroding the image utilizing disk-shaped structuring elements with radii of 5 and 3 pixels, respectively.

Then, we substitute the grayscale values of the pixels in the subtracted image, which correspond to vessels (attained in the registration step) with corresponding values in the auxiliary image (Fig. 6d). Thus, only vessels overlying areas of leakage are filled, without reducing the specificity of the algorithm by filling other dark area such as FAZ. We remove the remaining small outlier objects by applying an opening morphologic operator utilizing a disk-shaped structuring element with a radius of 2 pixels (Fig. 6e).

Leakage Segmentation. We deem all pixels with positive gray-level values in the resulting image as pilot estimates of the leakage area. We then utilize the contour of these pilot leakage regions to initialize Chan-Vese's active contour segmentation algorithm.[42] We empirically chose the parameters of the Chan-Vese algorithm (500 iterations and 0.8 for the smoothing parameter).

Detection of the Region of Interest (ROI) for Quantitative Analysis. We focused our quantitative analysis on a 1500-μm-radius circle around the fovea, which is of most significance for clinical diagnosis and treatment. Automatic designation of this region required detection of the fovea. Foveal identification on FA, regardless of utilization of automatic or manual methods, is a challenging problem especially in noisy real-world clinical data. We have developed an objective automatic algorithm to segment the fovea based on early FA frames, which are less affected by capillary nonperfusion and leakage as compared to later frames. We utilized this objective method only to determine the ROI for quantitative comparison of manual versus automatic grading. Indeed, better estimates for the center of the fovea can be attained by using alternative imaging modalities such as OCT.

Our automatic detection of fovea based on early FA frames was accomplished in the following steps: (1) applying an opening operation utilizing a disk-shaped structuring element with a radius of 50 pixels and (2) attaining the location of the fovea by averaging the coordinates of the darkest pixels in the central region of the image (defined as pixels with gray-level values less than 0.04 of the maximum intensity pixel in the region). Figure 6f illustrates the final extracted leakage area after applying the Chan-Vese algorithm on Figure 6e.
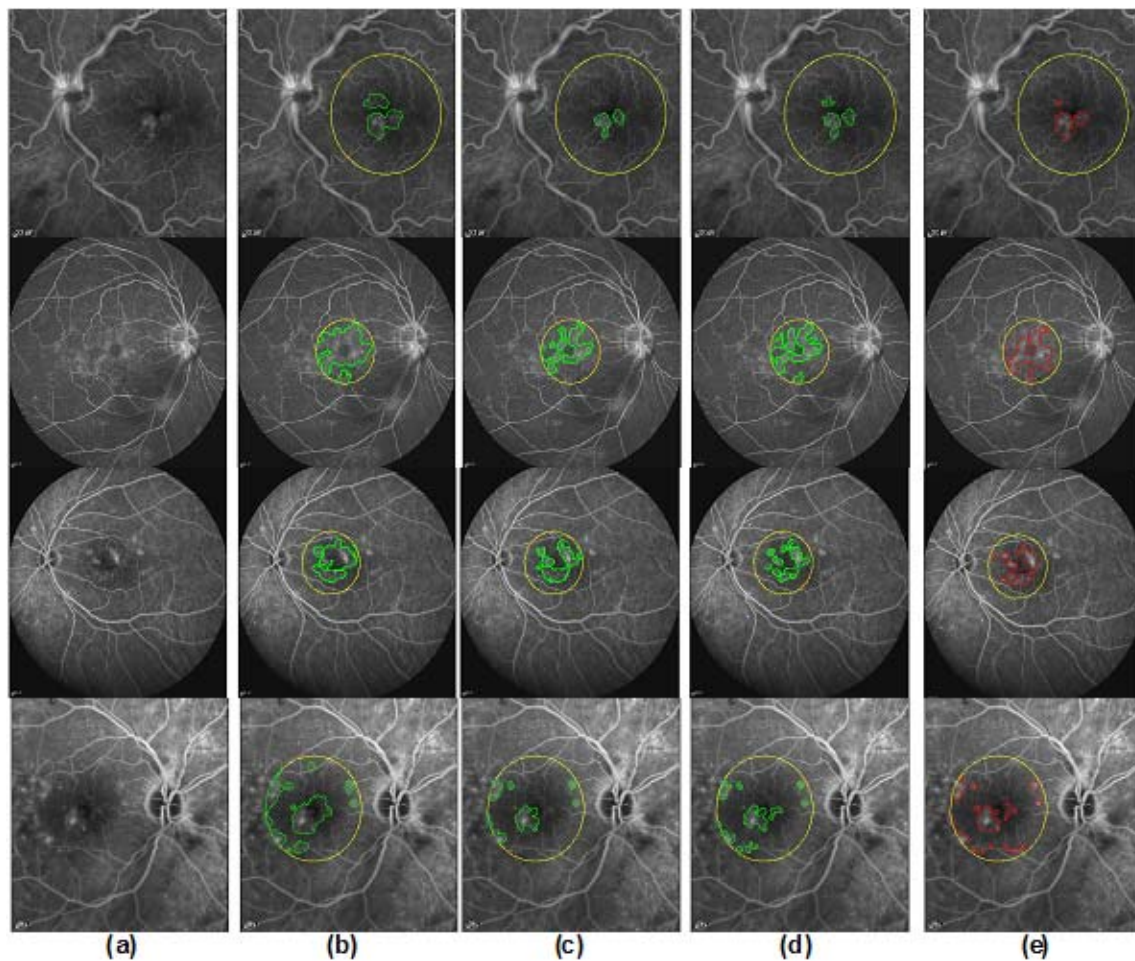
FIGURE 7. Comparison of leakage segmentation by manual graders (*green labels*) and automated method (*red labels*) in the ROI marked by the 3000-μm-diameter *yellow circle* centered at the fovea. (**a**) Late FA frame. (**b**) Segmented leakage by grader 1. (**c**) Segmented leakage by grader 2. (**d**) Resegmented leakage by grader 2 (at least 6 weeks later). (**e**) Segmented leakage by our algorithm. The FA videos in the *first and fourth rows* were captured at 30° FOV while the FA videos in the *second and third rows* were captured at 55° FOV.

Manual Segmentation. Total leakage was segmented in the late-phase FA images by two independent expert graders (MJA and PSM, both expert medical retina specialists) using the DOCTRAP software.[63] DOCTRAP has a graphic user interface (GUI) for manual segmentation extensively used and validated in previous studies.[64] Before commencing to grade the test dataset, manual graders met and agreed upon similar leakage definition and segmentation protocol defined by the senior clinician (SWC). To define intraobserver reliability, one manual grader repeated his grading on the same images at least 6 weeks after the initial grading. While grading, both early- and late-phase FA images were available to the reviewers on separate computer screens. Graders identified leakage as increased hyperfluorescence above the general choroidal background level present in the late but not the early phase. Early hyperfluorescent structures that did not leak, such as staining laser scars and nonleaking MAs, were not segmented as leakage. Similarly, preretinal neovascularization, identified as early bright hyperfluorescence with extensive, bright late leakage, was not considered leakage due to DME.

Quantitative Measures of Performance. In order to evaluate the performance of our algorithm, we calculated the specificity and sensitivity as follows. True positive (TP) was defined as the common segmented area (the number of

corresponding pixels in the ROI) by both the algorithm and the ophthalmologist. False positive (FP) was defined as an automatically segmented leakage area that does not belong to the leakage region as determined by the ophthalmologist. True negative (TN) is the area that does not belong to the detected leakage areas as determined by both the ophthalmologist and our algorithm. False negative (FN) is the area that was marked as a leakage region by the ophthalmologist but was missed by our algorithm. Sensitivity (TP/[TP+FN]), specificity (TN/[TN+FP]), and accuracy ([TP+TN]/[TP+TN+FP+FN]) for all data were calculated and compared to inter- and intraobserver errors.

Reproducibility Analysis. To test the reproducibility of the proposed algorithm, we divided each FA sequence into two separate sequences. One sequence included only the odd-numbered frames and the other included only the even-numbered frames of the original sequence. We compared the performance of the automatic algorithm in segmenting leakage area in these two sets of images from the same patient.

RESULTS

Figure 7 qualitatively compares the performance of our algorithm to the segmentation of manual graders. Table 2 lists the sensitivity, specificity, and accuracy of the automatic and manual grading for all datasets. The interobserver columns

TABLE 2. Quantitative Analysis of the Performance of the Proposed Automated Segmentation and Manual Grading of the Leakage Area in FA Images

Data	Automatic vs. Manual			Manual Interobserver			Manual Intraobserver		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
Diffuse 1	0.91	0.98	0.98	0.99	0.95	0.95	0.94	0.99	0.99
Diffuse 2	0.96	0.89	0.96	0.96	0.89	0.95	0.71	0.99	0.73
Diffuse 3	0.87	0.93	0.93	0.85	0.96	0.96	0.71	0.99	0.98
Diffuse 4	0.91	0.60	0.67	0.93	0.56	0.65	0.87	0.84	0.85
Diffuse 5	0.51	0.97	0.80	0.88	0.88	0.88	0.76	0.97	0.89
Diffuse 6	0.79	0.77	0.79	0.98	0.12	0.77	0.77	0.75	0.77
Diffuse 7	0.70	0.80	0.74	0.99	0.08	0.63	0.78	0.87	0.81
Focal 1	0.62	0.95	0.87	0.87	0.79	0.81	0.65	0.97	0.90
Focal 2	0.60	0.97	0.96	0.99	0.85	0.85	0.88	1	0.99
Focal 3	0.73	0.77	0.77	0.93	0.91	0.91	0.64	0.99	0.96
Focal 4	0.55	0.92	0.75	0.89	0.81	0.85	0.61	0.97	0.81
Focal 5	0.35	0.99	0.97	0.95	0.95	0.95	0.78	0.99	0.99
Focal 6	0.77	0.88	0.87	0.97	0.90	0.91	0.77	0.99	0.97
Focal 7	0.82	0.91	0.90	0.95	0.92	0.92	0.64	1	0.98
Focal 8	0.82	0.98	0.97	0.98	0.94	0.95	0.88	0.98	0.97
Focal 9	0.62	0.95	0.89	0.98	0.70	0.74	0.82	0.93	0.91
Focal 10	0.66	0.97	0.94	0.94	0.93	0.93	0.85	0.96	0.95
Mixed 1	0.70	0.95	0.82	0.84	0.69	0.76	0.73	0.74	0.74
Mixed 2	0.39	0.87	0.85	0.91	0.97	0.96	0.76	0.99	0.98
Mixed 3	0.80	0.98	0.90	1	0.59	0.78	0.92	0.97	0.95
Mixed 4	0.78	0.95	0.93	0.99	0.86	0.87	0.79	0.98	0.96
Mixed 5	0.56	0.92	0.82	0.99	0.55	0.68	0.81	0.85	0.84
Mixed 6	0.56	0.95	0.81	0.98	0.38	0.59	0.87	0.86	0.86
Mixed 7	0.67	0.97	0.80	0.97	0.27	0.58	0.86	0.89	0.88
Mean ± SD	0.69 ± 0.16	0.91 ± 0.09	0.86 ± 0.08	0.95 ± 0.05	0.73 ± 0.27	0.83 ± 0.16	0.78 ± 0.09	0.94 ± 0.08	0.90 ± 0.08

SD, standard deviation.

compare the performance of the two manual graders, while the intraobserver columns compare the performance of the same grader at two different time points at least 6 weeks apart. In our dataset, two subjects had evidence of prior macular photocoagulation (laser), five subjects had enlarged or irregular FAZs, two subjects had extrafoveal nonperfusion within the ROI, and seven subjects had definite foci of hemorrhage within the ROI. The mean area of leakage was 2.29 mm² in the ROI.

Note that no algorithmic parameter in our method was optimized based on the dataset that was used in our quantitative comparison.

According to Table 2, the mean accuracy was 0.86 ± 0.08 for automatic versus manual, 0.83 ± 0.16 for manual interobserver, and 0.90 ± 0.08 for manual intraobserver segmentation methods. To be more specific, the (sensitivity, specificity) of automatic versus manual grading for matching 30°, 35°, and 55° FOVs were (0.60, 0.88), (0.60, 0.97), and (0.73, 0.91), respectively. The (sensitivity, specificity) of manual interobserver grading for matching 30°, 35°, and 55° FOVs were (0.97, 0.58), (0.99, 0.85), and (0.94, 0.75), respectively. The (sensitivity, specificity) of manual intraobserver grading for matching 30°, 35°, and 55° FOVs were (0.80, 0.91), (0.88, 1), and (0.80, 0.92), respectively.

The reproducibility of the proposed algorithm in terms of accuracy, sensitivity, and specificity of detected leakage by our algorithm on average was 0.0034 ± 0.012, 0.0367 ± 0.0393, and 0.0152 ± 0.0239 pixels, respectively.

To facilitate comparison and future studies by other groups, we have made all the images used in the study (including raw FA videos and composite images) and their corresponding manual and automatic segmentation available at http://www.duke.edu/~sf59/Rabbani_IOVS_2014_dataset.htm.

DISCUSSION

We have presented a novel fully automatic algorithm for segmentation of leakage area on real-world clinical FA images, which was congruent with expert manual segmentation. Noting the quantitative results of Table 2, illustrated visually in Figure 7, although both graders followed the same protocol in identifying leakage, it is noteworthy that the interobserver accuracy was lower than for our automatic method. Moreover, the accuracy of our algorithm was close to the intraobserver accuracy (one grader versus himself), which is the highest practical value for accuracy (it is meaningless for an automatic algorithm to have higher accuracy than the gold standard of human grading, to which it is being compared). These results were achieved despite the fact that our (non-"cherry-picked") dataset suffered from noise and other distortions common in real-world clinical imaging. Figure 8 shows that in these situations, even intraobserver accuracy decreased greatly. We used the exact same algorithmic parameters for all experiments even though there was significant difference between imaging conditions (e.g., FOV) of different subjects. Indeed, we expect that we could have achieved better performance if we had selected images from a strict imaging protocol. However, our goal was to develop an algorithm that is useful for real-world clinical data, which are often far from the ideal situations considered in some clinical trials.

The main limitation of our algorithm is its inaccuracy in segmentation of relatively small leakage areas (e.g., Focal 5 and Mixed 2), resulting in lower reported sensitivity in subjects with relatively small leakage areas. However, as expected, the specificity values for these subjects are equal to if not better than the average specificity values across all subjects.

Another problem, which can be solved using high-speed computers, is the computational time of our algorithm due to registration of frames (which is around 2 minutes using
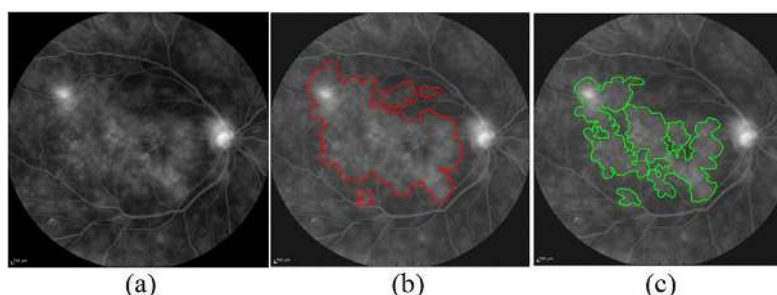
FIGURE 8. An example of the intraobserver reliability experiment in which an expert grader manually segmented the same image at two different time points. (**a**) A sample late FA image. (**b**) Manual segmentation of leakage by the expert grader at baseline. (**c**) Manual resegmented leakage area of the same image by the same grader after 6 weeks.

MATLAB R2013b for 8-bit 512×512 grayscale frames on a desktop PC with an Intel Core i7-4770 CPU @ 3.40 GHz, 8 GB RAM, 64-bit Windows 7 OS). Of course, this issue can be addressed in a commercial setting by coding this method for a graphics processing unit (GPU).

We also note that despite the robustness of our method to various sources of outliers, naturally the performance of our algorithm is negatively affected when dealing with significantly lower signal-to-noise ratio images. As part of our future work, to improve the signal-to-noise ratio of captured images, we will adapt a novel sparsity-based image enhancement algorithm, which has already demonstrated to be effective in enhancement of OCT images.[65,66]

Although FA provides additional information about DME that is complementary to OCT, change in leakage in FA is considered by many to be a more valuable metric than the absolute leakage at a single time point. This is in part because quantification of features on FA is typically not as reproducible compared to other imaging modalities such as OCT. The current study can be considered the first step toward automatic quantification of change in leakage over time.

Although several studies have been performed on quantitative analysis of various pathologies in FA images[10-12,67] (and other modalities including color fundus images[68-70] and OCT[63,71]), only a few papers have addressed automatic leakage detection for DME using FA.[10,34-39] Robust segmentation of leakage in clinical-grade data is a very difficult proposition, in part because of the challenging problem of FA sequence registration. This registration problem is challenging because (1) the deformation model is nonrigid, (2) the intensity of the images both locally and globally changes through time, (3) different sources of outliers locally (e.g., eye lashes) and globally (eye blinking) occlude the FOV, (4) the dynamic scene changes (e.g., leakage appears in the later frames). While each of these problems individually has been addressed in literature, a unique feature of our algorithm is its capability to fully automatically register FA images in the presence of outliers and significant leakage.

Attainment of an appropriate dataset for evaluating the reproducibility of our leakage detection algorithm was a challenging problem. Because FA imaging is invasive, repeated injection of fluorescein dye for research purposes was not permitted by the IRB. Moreover, even if repeated imaging of subjects was possible, repeatability in FA imaging is more an issue of variability in the imaging condition at two different time points (e.g., angle of incident of the laser) than of the robustness of the segmentation technique. To address this issue, we divided the images from the same imaging session into two nonoverlapping groups to demonstrate the repeatability of the algorithm without significant variability in imaging conditions.

In summary, here we introduce a new algorithm for automatic quantification of leakage in FA images of DME patients. The algorithm was based on nonrigid registration of FA frames, producing mean early FA and late FA images, obtaining the difference image, vessel filling and postprocessing, thresholding for obtaining the initial contour of the active contour, and leakage extraction in ROI using the Chan-Vese algorithm. While some of the algorithmic steps developed here were previously described by others, the overall algorithm is unique and novel, and shows unparalleled performance for segmenting leakage from real-world clinical FA images. This algorithm is implemented as MATLAB-based, user-friendly software, which has the potential to replace or aid subjective and time-consuming manual segmentation. Evaluation of usability and validation of this software for automatic classification of DME patients into focal, diffuse, and mixed categories in a clinical trial is part of our ongoing work. This novel, computer-aided technology will ultimately help us better understand the underlying mechanisms of diabetic retinopathy, which in turn may facilitate the optimal therapeutic strategy personalized for an individual's particular DME disease.

References

1. Centers for Disease Control and Prevention. National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2011.

2. Yau JW, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*. 2012; 35:556–564.

3. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. The Wisconsin epidemiologic study of diabetic retinopathy: IV. Diabetic macular edema. *Ophthalmology*. 1984;91:1464–1474.

4. Marmor MF. Mechanisms of fluid accumulation in retinal edema. *Doc Ophthalmol*. 1999;97:239–249.

5. Bringmann A, Pannicke T, Grosche J, et al. Muller cells in the healthy and diseased retina. *Prog Retin Eye Res*. 2006;25:397–424.

6. Schmidt-Erfurth U, Lang GE, Holz FG, et al. Three-year outcomes of individualized ranibizumab treatment in patients with diabetic macular edema: the RESTORE extension study. *Ophthalmology*. 2014;121:1045–1053.

7. Nguyen QD, Brown DM, Marcus DM, et al. Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. *Ophthalmology*. 2012;119:789–801.

8. Browning DJ, Glassman AR, Aiello LP, et al. Optical coherence tomography measurements and analysis methods in optical coherence tomography studies of diabetic macular edema. *Ophthalmology*. 2008;115:1366–1371, e1361.

9. Bhagat N, Grigorian RA, Tutela A, Zarbin MA. Diabetic macular edema: pathogenesis and treatment. *Surv Ophthalmol*. 2009; 54:1–32.

10. Smith RT, Lee CM, Charles HC, Farber M, Cunha-Vaz JG. Quantification of diabetic macular edema. *Arch Ophthalmol*. 1987;105:218–222.

11. Arzabe C, Jalkh A, Fariza E, Akiba J, Quiroz M. A simple device to standardize measurements of retinal structures in fundus photographs and retinal angiograms. *Am J Ophthalmol*. 1990; 109:107–108.

12. Barthes A, Conrath J, Rasigni M, Adel M, Petrakian J-P. Mathematical morphology in computerized analysis of angiograms in age-related macular degeneration. *Med Phys*. 2001; 28:2410–2419.

13. Hipwell J, Manivannan A, Vieira P, Sharp P, Forrester J. Quantifying changes in retinal circulation: the generation of parametric images from fluorescein angiograms. *Physiol Meas*. 1998;19:165–180.

14. Ibañéz MV, Simó A. Bayesian detection of the fovea in eye fundus angiographies. *Pattern Recognit Lett*. 1999;20:229–240.

15. Landini G, Misson GP, Murray PI. Fractal analysis of the normal human retinal fluorescein angiogram. *Curr Eye Res*. 1993;12: 23–27.

16. Landini G, Murray PI, Misson GP. Local connected fractal dimensions and lacunarity analyses of 60 degrees fluorescein angiograms. *Invest Ophthalmol Vis Sci*. 1995;36:2749–2755.

17. Chakravarthy U, Walsh AC, Muldrew A, Updike PG, Barbour T, Sadda SR. Quantitative fluorescein angiographic analysis of choroidal neovascular membranes: validation and correlation with visual function. *Invest Ophthalmol Vis Sci*. 2007;48:349–354.

18. Koprowski R, Teper S, Weglarz B, Wylegala E, Krejca M, Wróbel Z. Fully automatic algorithm for the analysis of vessels in the angiographic image of the eye fundus. *Biomed Eng Online*. 2012;11:35.

19. Kanagasingam Y, Bhuiyan A, Abràmoff MD, Smith RT, Goldschmidt L, Wong TY. Progress on retinal image analysis for age related macular degeneration. *Prog Retin Eye Res*. 2014;38: 20–42.

20. Zhou L, Rzeszotarski MS, Singerman LJ, Chokreff JM. The detection and quantification of retinopathy using digital angiograms. *IEEE Trans Med Imaging*. 1994;13:619–626.

21. Jagoe R, Blauth CI, Smith PL, Arnold JV, Taylor K, Wootton R. Automatic geometrical registration of fluorescein retinal angiograms. *Comput Biomed Res*. 1990;23:403–409.

22. Baudoin C, Lay B, Klein J. Automatic detection of microaneurysms in diabetic fluorescein angiography. *Rev Epidemiol Sante Publique*. 1983;32:254–261.

23. Frame AJ, Undrill PE, Cree MJ, et al. A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms. *Comput Biol Med*. 1998;28:225–238.

24. Cree MJ, Olson JA, McHardy KC, Sharp PF, Forrester JV. A fully automated comparative microaneurysm digital detection system. *Eye*. 1997;11:622–628.

25. Spencer T, Olson JA, McHardy KC, Sharp PF, Forrester JV. An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. *Comput Biomed Res*. 1996;29:284–302.

26. Spencer T, Phillips RP, Sharp PF, Forrester JV. Automated detection and quantification of microaneurysms in fluorescein angiograms. *Graefes Arch Clin Exp Ophthalmol*. 1992;230: 36–41.

27. Alipour SHM, Rabbani H. Automatic detection of microaneurysms in retinal images based on curvelet transform and morphological operations. *SPIE Optical Engineering + Applications*. 2013;8856:88561W.

28. Soltanipour A, Sadri S, Rabbani H, Akhlaghi M, Doost-Hosseini A. Vessel centerlines extraction from fundus fluorescein angiogram based on Hessian analysis of directional curvelet subbands. In: *Proc IEEE 2013 Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 2013:1070–1074.

29. Conrath J, Giorgi R, Raccah D, Ridings B. Foveal avascular zone in diabetic retinopathy: quantitative vs qualitative assessment. *Eye*. 2005;19:322–326.

30. Conrath J, Valat O, Giorgi R, et al. Semi-automated detection of the foveal avascular zone in fluorescein angiograms in diabetes mellitus. *Clin Experiment Ophthalmol*. 2006;34:119–123.

31. Haddouche A, Adel M, Rasigni M, Conrath J, Bourennane S. Detection of the foveal avascular zone on retinal angiograms using Markov random fields. *Digit Signal Process*. 2010;20: 149–154.

32. Zheng Y, Gandhi JS, Stangos AN, Campa C, Broadbent DM, Harding SP. Automated segmentation of foveal avascular zone in fundus fluorescein angiography. *Invest Ophthalmol Vis Sci*. 2010;51:3653–3659.

33. Alipour SHM, Rabbani H, Akhlaghi M. A new combined method based on curvelet transform and morphological operators for automatic detection of foveal avascular zone. *Signal Image Video Process*. 2014;8:205–222.

34. Phillips R, Spencer T, Ross P, Sharp P, Forrester J. Quantification of diabetic maculopathy by digital imaging of the fundus. *Eye*. 1991;5:130–137.

35. Martinez-Costa L, Marco P, Ayala G, De Ves E, Domingo J, Simó A. Macular edema computer-aided evaluation in ocular vein occlusions. *Comput Biomed Res*. 1998;31:374–384.

36. Cree MJ, Olson JA, McHardy KC, Sharp PF, Forrester JV. The preprocessing of retinal images for the detection of fluorescein leakage. *Phys Med Biol*. 1999;44:293–308.

37. Phillips RP, Ross PG, Tyska M, Sharp PF, Forrester JV. Detection and quantification of hyperfluorescent leakage by computer analysis of fundus fluorescein angiograms. *Graefes Arch Clin Exp Ophthalmol*. 1991;229:329–335.

38. Phillips R, Ross P, Sharp P, Forrester J. Use of temporal information to quantify vascular leakage in fluorescein angiography of the retina. *Clin Phys Physiol Meas*. 1990;11:81–85.

39. Buchanan CR, Trucco E. Contextual detection of diabetic pathology in wide-field retinal angiograms. *Conf Proc IEEE Eng Med Biol Soc*. 2008;5437–5440.

40. Berger JW. Quantitative spatiotemporal image analysis of fluorescein angiography in age-related macular degeneration. *BiOS'98 International Biomedical Optics Symposium: International Society for Optics and Photonics*. 1998:48–53.

41. El-Shahawy MS, ElAntably A, Fawzy N, Samir K, Hunter M, Fahmy AS. Segmentation of diabetic macular edema in fluorescein angiograms. In: *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 2011: 661–664.

42. Chan TF, Vese LA. Active contours without edges. *IEEE Trans Image Process*. 2001;10:266–277.

43. Coscas-Mimoun FMG, Kone T, Bunel P, Coscas G. Completely automatic overlay of retinal fluorescein angiographic pictures. *Invest Ophthalmol Vis Sci*. 1992;4:732.

44. Dréo J, Nunes J-C, Siarry P. Robust rigid registration of retinal angiograms through optimization. *Comput Med Imaging Graph*. 2006;30:453–463.

45. Choe TE, Cohen I. Registration of multimodal fluorescein images sequence of the retina. In: *ICCV 2005 Tenth IEEE International Conference on Computer Vision*. 2005:106–113.

46. Domingo J, Ayala G, Simó A, de Ves E, Martínez-Costa L, Marco P. Irregular motion recovery in fluorescein angiograms. *Pattern Recognit Lett*. 1997;18:805–821.

47. Nunes JC, Bouaoune Y, Delechelle E, Bunel P. A multiscale elastic registration scheme for retinal angiograms. *Comput Vis Image Underst*. 2004;95:129–149.

48. Kubecka L, Jan J, Kolar R, Jirik R. Elastic registration for autofluorescence image averaging. *Conf Proc IEEE Eng Med Biol Soc*, 2006:1948–1951.

49. Tsai C-L, Li C-Y, Yang G, Lin K-S. The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence. *IEEE Trans Med Imaging*. 2010;29:636–649.

50. Stewart CV, Tsai C-L, Roysam B. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Trans Med Imaging*. 2003;22:1379–1394.

51. Perez-Rovira A, Cabido R, Trucco E, McKenna SJ, Hubschman JP. RERBEE: Robust Efficient Registration via Bifurcations and Elongated Elements applied to retinal fluorescein angiogram sequences. *IEEE Trans Med Imaging*. 2012;31:140–150.

52. Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. 1981;24:381–395.

53. Torr PH, Zisserman A. MLESAC: a new robust estimator with application to estimating image geometry. *Comput Vision Image Underst*. 2000;78:138–156.

54. Estrada R, Tomasi C, Cabrera MT, Wallace DK, Freedman SF, Farsiu S. Exploratory Dijkstra forest based automatic vessel segmentation: applications in video indirect ophthalmoscopy (VIO). *Biomed Opt Express*. 2012;3:327–339.

55. Esmaeili M, Rabbani H, Mehri A, Dehghani A. Extraction of retinal blood vessels by curvelet transform. *2009 16th IEEE International Conference on Image Processing (ICIP)*. 2009: 3353–3356.

56. Niemeijer M, Staal J, van Ginneken B, Loog M, Abramoff MD. Comparative study of retinal vessel segmentation methods on a new publicly available database. *Medical Imaging 2004*. 2004:648–656.

57. Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). *Comput Vis Image Underst*. 2008;110:346–359.

58. Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. *IEEE Trans Pattern Anal Mach Intell*. 2005;27: 1615–1630.

59. Alahi A, Ortiz R, Vandergheynst P. Freak: fast retina keypoint. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2012:510–517.

60. Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comput Vis*. 2004;60:91–110.

61. Gonzalez RC, Woods RE. *Digital Image Processing*. 3rd ed. Upper Saddle River, NJ: Prentice Hall; 2008:665–679.

62. Pizer SM, Amburn EP, Austin JD, et al. Adaptive histogram equalization and its variations. *Comput Vis Graph Image Process*. 1987;39:355–368.

63. Chiu SJ, Izatt JA, O'Connell RV, Winter KP, Toth CA, Farsiu S. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. *Invest Ophthalmol Vis Sci*. 2012;53:53–61.

64. Lee JY, Chiu SJ, Srinivasan PP, et al. Fully automatic software for retinal thickness in eyes with diabetic macular edema from images acquired by Cirrus and Spectralis systems. *Invest Ophthalmol Vis Sci*. 2013;54:7595–7602.

65. Fang L, Li SH, McNabb RP, et al. Fast acquisition and reconstruction of optical coherence tomography images via sparse representation. *IEEE Trans Med Imaging*. 2013;32: 2034–2049.

66. Kafieh R, Rabbani H, Selesnick IW. Three dimensional data-driven multi scale atomic representation of optical coherence tomography. *IEEE Trans Med Imaging*. In press.

67. Friedman D, Parker JS, Kimble JA, Delori FC, McGwin G Jr, Curcio CA. Quantification of fluorescein-stained drusen associated with age-related macular degeneration. *Retina*. 2012;32:19–24.

68. Smith R, Chan J, Nagasaki T, Sparrow J, Barbazetto I. A method of drusen measurement based on reconstruction of fundus background reflectance. *Br J Ophthalmol*. 2005;89:87–91.

69. Smith RT, Chan JK, Nagasaki T, et al. Automated detection of macular drusen using geometric background leveling and threshold selection. *Arch Ophthalmol*. 2005;123:200–206.

70. Sohrab MA, Smith RT, Salehi-Had H, Sadda SR, Fawzi AA. Image registration and multimodal imaging of reticular pseudodrusen. *Invest Ophthalmol Vis Sci*. 2011;52:5743–5748.

71. Farsiu S, Chiu SJ, O'Connell RV, et al. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. *Ophthalmology*. 2014;121:162–172.