33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Statistical Characterization and Segmentation of Drusen in Fundus Images

H. Santos-Villalobos, *Member, IEEE*, T. P. Karnowski, *Member, IEEE*, D. Aykac, *Member, IEEE*, L. Giancardo, *Member, IEEE*, Y. Li, *Member, IEEE*, T. Nichols, M.D., *Member, IEEE*, K.W. Tobin, Jr., *Senior Member, IEEE*, E. Chaum, M.D., *Member, IEEE*

***Abstract*— Age related Macular Degeneration (AMD) is a disease of the retina associated with aging. AMD progression in patients is characterized by drusen, pigmentation changes, and geographic atrophy, which can be seen using fundus imagery. The level of AMD is characterized by standard scaling methods, which can be somewhat subjective in practice. In this work we propose a statistical image processing approach to segment drusen with the ultimate goal of characterizing the AMD progression in a data set of longitudinal images. The method characterizes retinal structures with a statistical model of the colors in the retina image. When comparing the segmentation results of the method between longitudinal images with known AMD progression and those without, the method detects progression in our longitudinal data set with an area under the receiver operating characteristics curve of 0.99.**

I. INTRODUCTION

In the United States, more than 25 million Americans are estimated to have diabetes, and the number of adults with the disease is projected to exceed 110 million by the year 2050 [1]. Worldwide, diabetes is also a growing problem and the need for broad-based inexpensive screening of the eyes of diabetics cannot be met by the current practice of screening by ophthalmologists. Automated retinal image processing research thus has become a leading topic of research in recent years as a possible means of combating this public health problem [2], [3], [4], [5]. However, other retina diseases are also benefiting from research in this area. In this work we consider age-related macular degeneration (AMD) [6] and the characterization of features of AMD through image segmentation. AMD is a chronic, progressive disease that is the leading cause of vision loss in aging patients. For the patient, the disease causes central vision blurring, distortion, and vision impairment, often to the level of legal blindness. There are two major types of AMD termed wet, or neovascular, and dry, or atrophic. The former consists of the growth of fragile new blood vessels, which often leak fluid, causing the macula to swell and leading to scarring and vision loss. In the more common dry AMD, the effects are more subtle as the cells of the sensory retina degenerate over time. The hallmark manifestation of AMD is the formation of soft drusen, fuzzy yellowish deposits beneath the retina, which may increase over time as the disease progresses. As dry AMD progresses, visible pigmentary changes in the fundus may occur as retinal pigment epithelial tissue beneath the retina, which is critical for normal retinal function, is lost. Later effects include Geographic Atrophy (GA), which is a distinct, nummular area of retina pigmentary atrophy. Two suggested standards for categorization of AMD with respect to retina features are summarized in Tables I and II [6]. The first, from the Age-Related Eye Disease Study (AREDS) uses a four-category scale for AMD, based on the presence or absence of key phenomena. The second assigns points, based on the presence of these features [7] in each eye, and the total points from both eyes are added. Scores of 1, 2, 3, and 4 are assigned AMD risks of 3%, 12%, 25%, and 50%.

TABLE I
AREDS AMD SCALE

Category	Description
1	Less than 5 small (i.e., $< 63\mu m$) drusen
2	Extensive small drusen, or at least one intermediate (i.e., $< 125\mu m$) sized or pigment abnormalities
3	Extensive intermediate sized drusen; at least one large (i.e., $> 125\mu m$) druse; non-central GA; or any combination of these
4	No advanced AMD (GA in center of fovea or choroidal neovascular membrane) in study eye, but advanced AMD in fellow eye

TABLE II
BRESSLER AMD SCALE

Phenomena	Points
One or more large drusen $> 125\mu m$	1
Multiple intermediate drusen and no large drusen	0.5
Pigmentation changes	1
Central geographic atrophy	2
Neovascular AMD	2

In the medical image processing literature, drusen detection is a subject of interest, although it generally has not received the attention of other pathological features of the retina such as microaneurysms (i.e., a dilation of a small retinal capillary vessel associated with diabetic retinopathy). Sbeh [8] used morphological reconstruction to detect drusen by identifying regions of relative maxima and minima, and

Manuscript received June 18th, 2011. This work was supported in part by the National Eye Institute of the National Institutes of Health (R01 EY017065), the Research to Prevent Blindness, New York, NY, the Plough Foundation, Memphis TN. This paper was prepared by OAK RIDGE NATIONAL LABORATORY, Oak Ridge, TN, USA, 37831-6285, operated by UT-BATTELLE, LLC for the US DEPARTMENT OF ENERGY under contract DE-AC05-00OR22725.

H. Santos-Villalobos, T.P. Karnowski, D. Aykac, L. Giancardo, T. Nichols, and K.W. Tobin, Jr., are with the Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA (e-mail: hsantos@ornl.gov)

E. Chaum and Y. Li are with the University of Tennessee Health Science Center, 930 Madison Avenue, Suite 731, Memphis TN 38163. Dr. Chaum is an RPB Senior Scientist.

performing morphological reconstruction around maxima, bounded by minima, and followed by a supervised learning approach to reject false positives. Brandon [9] used wavelet analysis on image regions along with rule-based classification to identify drusen. Rapantzikos [10] used a block-based method and thresholded each block with a level based on a classification of the region derived from the histogram statistics (skewness and kurtosis), or a further partition of the region into a finer group of blocks. Neimeijer [11] used a filter bank based on Gaussian derivatives to create features that generated a pixel-based probability of bright lession, which were clustered and then classified by a supervised learning algorithm to distinguish between drusen and other bright lesions.

In this work, we are concerned with segmentation of drusen, with the ultimate goal of characterizing the progression of drusen through a longitudinal image study. Overall our goal is to determine a more quantitative means of measuring the AMD state and progression using statistical and image processing methods. The paper is organized as follows. In the methodology section we present an overview of the processing used to characterize retinal structures (e.g., the retinal wall, and drusen lesions) in the retinal images. In the experimental section, we describe our data set and present the results of our analysis methods. We conclude with observations and projections for future research.

II. METHODOLOGY

In this section we discuss the key functional steps of the proposed algorithm including image normalization, training and drusen characterization, processing, and identification.

A. Image Normalization

We use a method based on [12] to normalize the fundus images to attempt to control for illumination changes. We assumed fundus image colors in the sRGB color space. The sRGB color space consists of three color channels red, green, and blue, discretized into the integer range [0 – 255]. First, the sRGB image is transformed to the CIE L*C*h* color space. The CIE L*C*h* color space represent colors in a sphere, where L* represents lightness, C* chroma, and h* the hue angle. We manipulate the L*-channel to reduce illumination variations without modifying the image's hue and chroma. A large median filter is applied to the L* color plane with a neighborhood size that is roughly 0.04% of the largest image side (i.e., 54×54 for a 1360×1024 pixel image). This median filtered image is subtracted from the original image to produce a difference image D, and then each pixel of D is normalized to standard score

$$D_N = \frac{D - \mu_D}{\sigma_D},$$

where μ_D and σ_D are the empirical mean and standard deviation of D, respectively. Then, the normalized pixels intensities in D_N are mapped to the L*-channel standard range [0 – 100]. The resulting CIE L*C*h* image is transformed back to the sRGB color space.

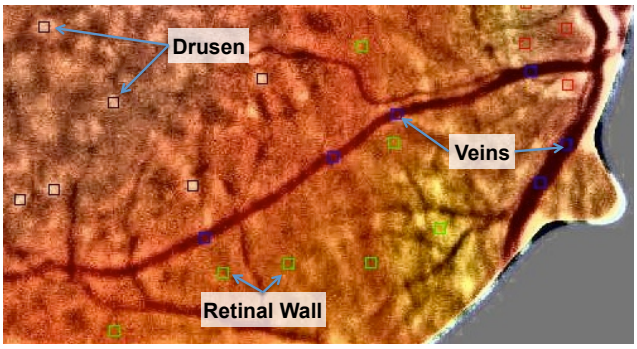


Fig. 1. Example of 12×12 window samples for different retinal structures.

B. Retina Characterization

Given that the eye is a natural phenomenon, we hypothesized that the color features of a retinal structure (e.g., a drusen lesion) may abide to an unknown conditional density function $\rho(\vec{X}|\omega_i)$, where $\vec{X}$ is a color feature vector in the CIE L*C*h* color space, and the class ω_i corresponds to a retinal structure. We claim that $\rho(\vec{X}|\omega_i)$ can be estimated by drawing samples of the class ω_i from normalized retinal images. Although it may be feasible to normalize retinal images from all ethnicities to the same scale, at the present, we apply our methodology to retinal images from Caucasian patients alone. We characterize two retinal structures by estimating $\rho(\vec{X}|WALL)$ and $\rho(\vec{X}|DRUSEN)$, the conditional density functions for the retinal wall and the drusen lesions, respectively. Next, we describe how these conditional density functions are estimated.

We used a 12×12 pixel window to collect samples from retinal images with a pixel size of roughly $6\mu m/pixel$ (See Fig. 1). The window's underlying pixels should belong to the same class ω_i. Consequently, each window contributes with 144 data samples for the corresponding retinal structure ω_i. The size of the window was selected with the purpose of making the window smaller than the retinal structures, while keeping a satisfactory sample size. We extracted samples for the following structures: blood vessels, retinal wall, and drusen lesions. The blood vessel samples were used to select the candidate drusen pixels, which will be discussed later. The samples were labeled and stored. Each data sample $\vec{X}$ consisted of three features, the lightness (L), saturation ($S = 100 \times C/\sqrt{C^2 + L^2}$), and hue ($h$) values of the corresponding CIE L*C*h* pixel. It is important to underscore that $\rho(\vec{X}|\omega_i)$ does not characterize the surface (i.e., texture) of the window sample, it rather models the likelihood that a pixels color triplet [L,S,h] belongs to the class ω_i. Although we use a window to collect samples, the spatial information is lost when the window is transformed to a column vector.

The steps to estimate $\rho(\vec{X}|\omega_i)$ are: 1) Collect all M 12×12 window samples that correspond to ω_i. 2) Transform the window samples to M 144-pixels arrays, and concatenate them to a single $M \times 144$-pixels array. 3) Individually, normalize each color channel to standard score. 4) Find the minimum, C_{min}, and maximum, C_{max}, values of the normalized values

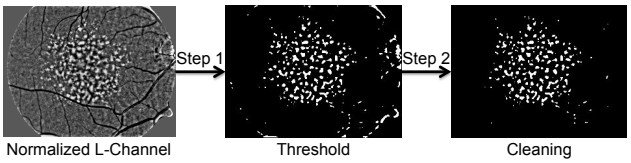


Fig. 2. Two-step pre-processing of retinal images to select drusen pixel candidates.

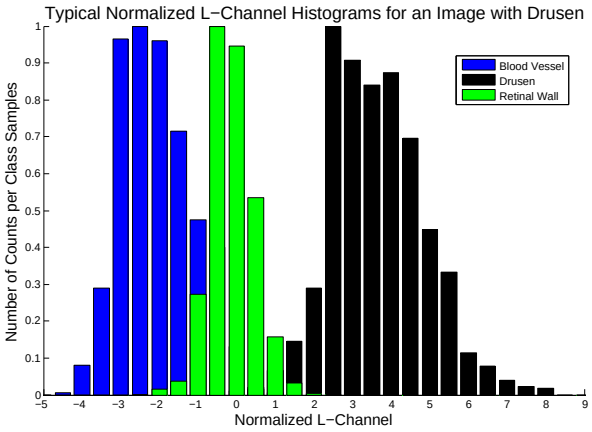


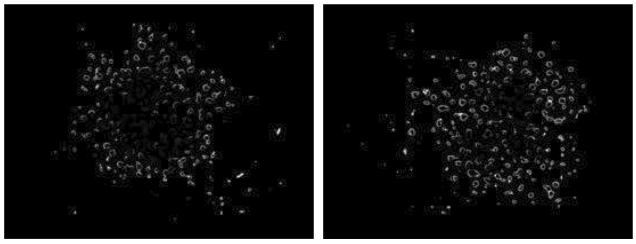
Fig. 3. Example of typical L-Channel histograms for retinal wall, blood vessels, and drusen lesion samples.

(All channel values together). 5) Create a three-dimensional histogram with range $[C_{min}, C_{max}]$. The number of bins depends on the number of available samples. We binned our training samples in 343 bins (7 bins per channel). A larger number of bins will spread the samples in the histogram, making the bin counts unreliable. 6) Bin the normalized samples in the three-dimensional histogram, and divide all the bin counts by the number of samples N ($N = M \times 144$). This last step will make $\int \rho(\vec{X}|\omega_i)d\vec{X} = 1$.

C. Drusen Pixel Candidates Selection

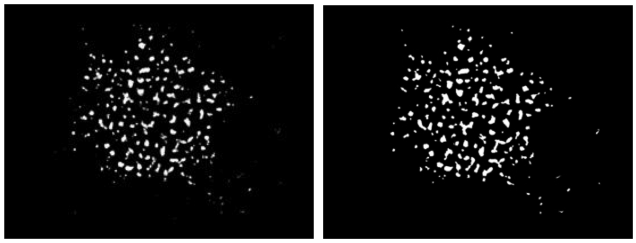
As explained above, our statistical segmentation of drusen works at the pixel level. Retinal images are composed of more than half a million pixels, which we would like to process in an efficient manner. Therefore, it is advantageous to remove from the segmentation process pixels that are unlikely to be part of a drusen lesion. To address this, the retinal image goes through a two-step pre-processing (See Fig. 2), with all operations performed in the normalized L*-channel. In Step 1, we compute a binary mask M_T with pixels "ON" when the corresponding image pixel intensities are greater than T and "OFF" otherwise. As shown in Fig. 3, our assessment of the training samples (obtained from three images of patients with significant drusen) showed that drusen pixel intensities are likely to be greater than one standard deviation over the mean of the image intensities; which translates to pixels intensities greater than one (i.e., T=1) for a normalized L*-channel. In Step 2, we turn "OFF" pixels from M_T that are close to the circular edge of the retinal image and the optical nerve. The retinal edge has a tendency to generate glare and bright artifacts; and the optical nerve usually appears as a bright disk in the retina. However, it is convenient to remove these structures from our candidate pixels, because both structures are far from the macula region. A vignette mask M_V is computed from for the retinal image and its border is dilated so the mask overlaps with the retinal edge. The center pixels of the mask are "ON" and the outside pixels are "OFF". Then, an additional mask M_{ON} is generated with a disk centered at the position of the optical nerve and with radius about 15% of the size of the image. The disc pixels are turned "OFF", and the mask remaining pixels are "ON". Finally, the drusen candidate pixel mask $M_{DC} = M_T \cap M_V \cap M_{ON}$. The remaining operations of the proposed algorithm are performed over these candidate pixels alone.

D. Drusen Identification



(a) (b)

Fig. 4. Likelihood that the drusen candidate pixels of the left image in Fig. 2 (a) belong to retinal wall, and (b) drusen lesion structures.



(a) (b)

Fig. 5. Final (a) likelihood ratio results, and (b) segmentation mask, M_D, for the drusen candidate pixels of the left image in Fig. 2

To classify the pixels as part of a retinal wall or a drusen lesion structure, we employ the likelihood ratio of the Neyman-Pearson lemma. First, from the estimated $\rho(\vec{X}|WALL)$ and $\rho(\vec{X}|DRUSEN)$, in Subsection II-B, we calculate the likelihood that each candidate pixel belongs to either class WALL or DRUSEN. Fig. 4 illustrates the likelihood results for both classes. Then, we can calculate the likelihood ratio

$$\log \frac{\rho(\vec{X}|DRUSEN)}{\rho(\vec{X}|WALL)} \underset{\leq}{>} \eta, \ \eta \in \mathbb{R}.$$

If the ratio is greater than η the pixel is classified as part of a drusen lesion, and as a retinal wall otherwise. Although an optimal $\eta*$ can be computed from training data, we fixed $\eta = 1$. This means that a pixel is classified as drusen when

$\rho(\vec{X}|DRUSEN) > \rho(\vec{X}|WALL)$. A binary mask M_D is created, with all pixels classified as drusen "ON" and the remaining pixels "OFF".

As a final step, morphological operations are performed to fill holes in the drusen mask, and to remove too small ($Area < 0.441mm^2$) or too large ($Area > 140.6mm^2$) drusen segmentations from M_D. These size limits are based on a fundus camera with pixel size of $6\mu m^2$. A final segmentation result is shown in Fig. 5.

III. EXPERIMENTS

In this section we describe the experimental results using data from the telemedicine network. Some background on the data set is provided, and we then describe our validation process. Finally we present our results on the analysis method.

A. Data Set

Three main data sets were used in this work. The first data set consists of AREDS2 data. From this set, we hand selected eyes from Caucasian patients. The AREDS2 subset has 42 patient eyes from 8 patients. One patient had only one eye imaged, and the rest had both eyes. In the set of images, all had at least two images taken roughly a year apart, 6 had three images, and 2 had two images. We used three of these images to compute the statistical properties of the retina structures as mentioned in Subsection II-D, at a rate of 12,816 (41.6%) retinal wall samples, 5,184 (16.8%) blood vessel samples, and 12,816 (41.6%) drusen samples. These training images were not used in the subsequent processing. A second set of images (TRIAD-A) was hand selected from the Telemedical Retina Imaging and Diagnoses (TRIAD) network [13]. These images consisted only of "true negatives", in that they were all diagnosed as normal retina with no significant lesions. Thus, any machine-segmentation images that were found on these images were regarded as true negatives. In this set, there were 7 images from 7 distinct patients.

Finally, a third set of images (TRIAD-B) was also taken from the TRIAD network, but these consisted of return patients to the system. These images served as examples of multiple patient visits where there was no noticeable disease or disease progression, as characterized by an ophthalmologist review. In this set, we further filtered by restricting the image quality as measured by the method of [14] of 0.8 or greater. In this set, there were 70 patient-eye examples, all from 45 degrees field of view from a total of 21 patients. In this set, one patient had 3 images from both eyes, another 12 patients had 2 images, and the rest had one per eye only.

B. Ground truth and Post-processing

Some post-processing was added to reduce false positives in normal images (See Fig. 6). The post-processing used a supervised learning algorithm with ground-truth data provided by reviewing images with known drusen. A morphological-based method based on [8] was used to tag potential drusen. The method has a tendency to over-segment which made it a good candidate for a ground-truth generator since we felt confident it would detect drusen well and could be manually filtered into true positive and false positive detections. Thus the ground-truth process consisted of hand-selected definitive, true-positive drusen candidates from the AREDS2 images. True negatives were identified by using the detections on images with no known drusen or other lesions. The candidate lesions detected by the statistical method were then grouped and a hold-one-out validation test was conducted. A set of features of each candidate lesion was computed [15] (e.g., simple shape-based features and color measurements). Images were held-out of the total data set on a per-patient basis, the neural network was trained on the ground-truth data, and the held-out images were filtered by classifying the candidates into 'true' lesions and 'nuisance' lesions. See Fig. 6 and Fig. 7 for drusen segmentation examples for a normal retina and abnormal retina, respectively.

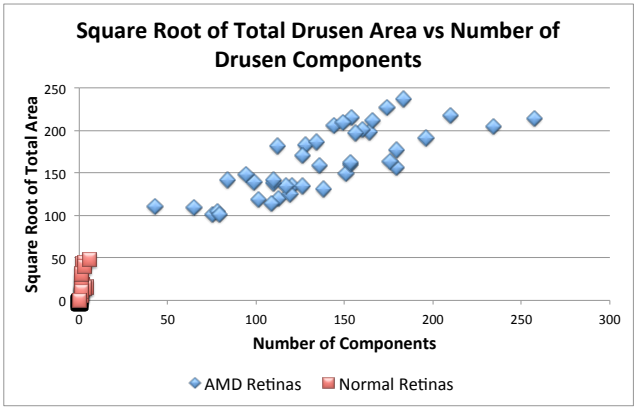
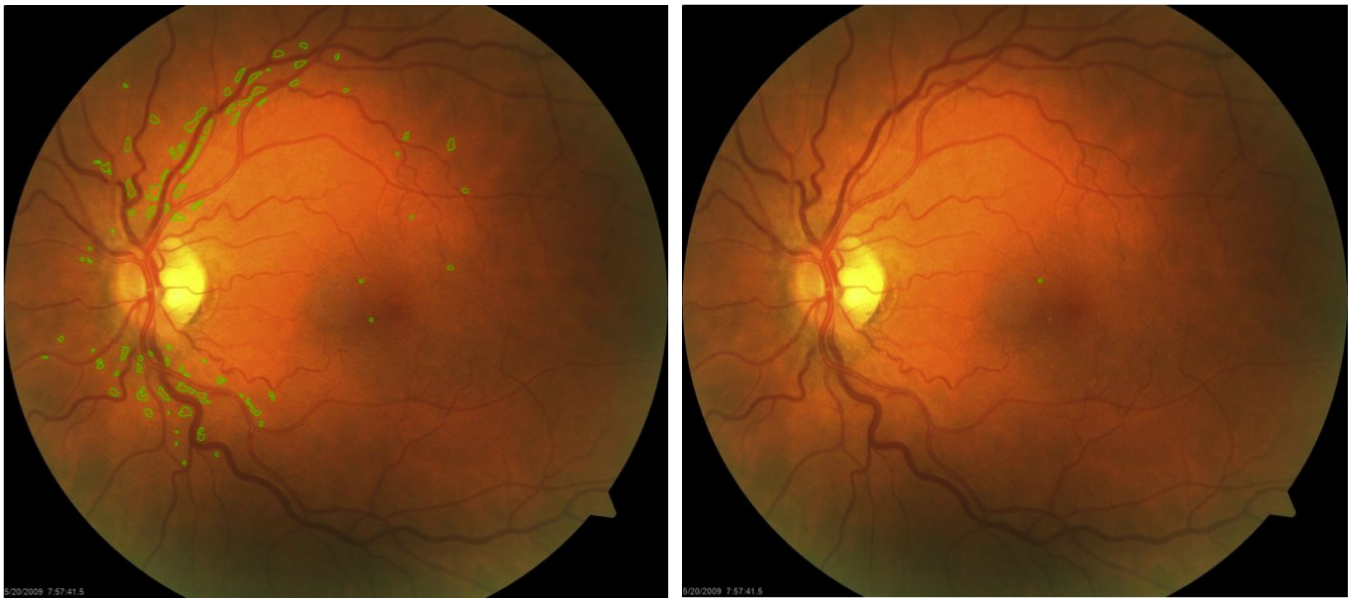


Fig. 8. Plot of detected drusen components and composite drusen area on a per-image basis. Observe the separation between normal and abnormal retinas for their total number of detections and square root total area.

C. Characterization results

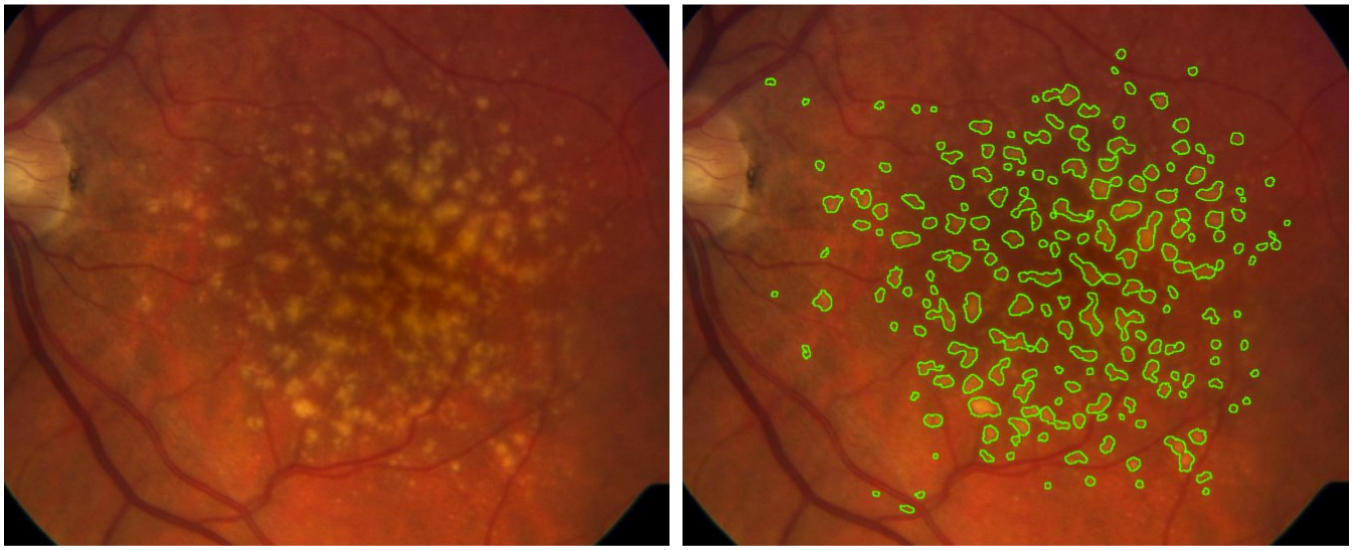
The neural network results revealed a sensitivity of 99% and a specificity of 98% averaged across all the images, but we note that these results apply only to the ground-truth data and may not accurately reflect true drusen detection since a completely missed lesion would not be accurately counted. After post-processing, the number of drusen detected and square-root of the total drusen area is shown in Fig. 8. This reveals very high separability between the images for this data set, with virtually perfect Receiver Operating Characteristic (ROC) curves. However, we again note that the data set is likely not as challenging since the AMD images in the set have a large number of drusen.

Another characterization of the dataset is the detection of changes over time in a patient eye. For these results, we show a ROC curve using the change in the number of detected drusen. See the ROC curves in Fig. 9. We again achieve excellent results (an area under curve of 0.99), but we would like to test the method on more challenging datasets.



(a) (b)

Fig. 6. Segmentation results for a patient with a normal retina. (a) Drusen detected by our statistical model. (b) Most false-positves removed by our post-processing algorithm.



(a) (b)

Fig. 7. Segmentation results for a patient with an abnormal retina. (a) Unprocessed original image. (b) Drusen detected by our statistical model. For this patient, the post-processing algorithm did not remove any components.

IV. CONCLUSIONS

We proposed a statistical model to characterize the retinal structures of fundus images and a methodology to detect drusen lesions. More precisely, we showed that an estimation of the conditional distribution of the color features in a retina image such as lightness, saturation, and hue, followed by a likelihood ratio classification, successfully detected the presence of drusen lesions in abnormal eyes. The method alone creates false positives in normal images. However, a post-processing of the drusen candidates removes these false positives when additional texture, shape, and spatial features are employed in the classification.

Further work in this area involves 1) increasing the sensitivity of the method, 2) improving our statistical model and extending it to characterize other retinal lesions, 3) extending the technique for retinas with different background pigmentation (e.g., retina images from African American, and Hispanic patients) 4) extending this study to larger data sets, in particular more challenging sets with more subtle progression or evidence of AMD, and 5) to use image analysis methods to establish a standard progression rating.

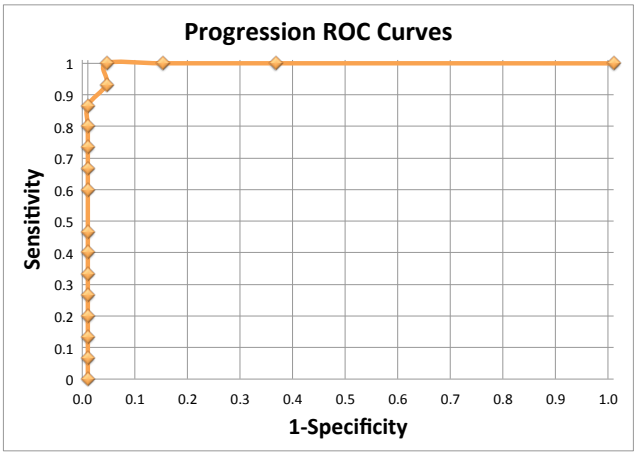


Fig. 9. ROC curve for detection of progression in longitudinal images. The change in number of drusen detected was used as the detection criteria. For this data set, the algorithm works very well at identifying image sets with significant change in drusen content.

We anticipate that when we extend this work to larger data sets, less heuristic choices for some of the algorithm parameters will become evident. Finally, the segmentation and characterization methods would be beneficial in a telemedical network such as TRIAD for characterizing imaged retinas and the progression method would be useful in detecting changes in returning patients.

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*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2011.

[2] M. D. Abramoff, M. Niemeijer, and S. R. Russell, "Automated detection of diabetic retinopathy: barriers to translation into clinical practice," *Expert Review Medical Devices*, vol. 7, no. 2, pp. 287–296, 2010.

[3] T. Teng, M. Lefley, and D. Claremont, "Progress towards automated diabetic ocular screening: A review of image analysis and intelligent systems for diabetic retinopathy," *Medical and Biological Engineering and Computing*, vol. 40, pp. 2–13, 2002.

[4] A. Fleming, K. Goatman, S. Philip, J. Olson, and P. Sharp, "Automatic detection of retinal anatomy to assist diabetic retinopathy screening," *Physics in Medicine and Biology*, vol. 52, no. 2, pp. 331–345, 2007.

[5] K. W. Tobin, M. D. Abramoff, E. Chaum, L. Giancardo, V. P. Govindasamy, T. P. Karnowski, M. T. Tennant, and S. Swainson, "Using a patient image archive to diagnose retinopathy," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, Vancouver, Canada, Aug. 2008, pp. 5441 –5444.

[6] D. Huang, P. Kaiser, C. Lowder, and E. Traboulsi, *Retinal Imaging*. Philadelphia, PA: Mosby Elsevier, 2005.

[7] S. Bressler, F. Ferris, M. Davis, R. Gangnon, L. Hubbard, L. Lee, E. Chew, B. Klein, R. Klein, and AREDS Research Group, "A simple clinical scale for estimating the risk of age-related macular degeneration progression," *Invest. Ophthalmol. Vis. Sci.*, vol. 46, no. 5, p. 2426, 2005.

[8] Z. Ben Sbeh, L. Cohen, G. Mimoun, and G. Coscas, "A new approach of geodesic reconstruction for drusen segmentation in eye fundus images," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 12, pp. 1321 –1333, Dec. 2001.

[9] L. Brandon and A. Hoover, "Drusen detection in a retinal image using multi-level analysis," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, R. Ellis and T. Peters, Eds., vol. 2878. Montreal, Canada: Springer Berlin / Heidelberg, Nov. 2003, pp. 618–625.

[10] K. Rapantzikos, M. Zervakis, and K. Balas, "Detection and segmentation of drusen deposits on human retina: potential in the diagnosis of age-related macular degeneration," *Medical Image Analysis*, vol. 7, no. 1, pp. 95–108, mar. 2003.

[11] M. Niemeijer, B. van Ginneken, S. Russell, M. Suttorp-Schulten, and M. Abramoff, "Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis," *Investigative Ophthalmology & Visual Science*, vol. 48, no. 5, pp. 2260–2267, 2007.

[12] M. Cree, E. Gamble, and D. Cornforth, "Colour normalisation to reduce inter-patient and intra-patient variability in microaneurysm detection in colour retinal images," in *WDIC 2005, APRS Workshop on digital image computing*, Brisbane, Australia, Feb. 2005, pp. 163–168.

[13] J. Spiegel, Y. Li, T. Karnowski, S. Morris, E. Sparrow, and E. Chaum, "A one-year outcomes assessment of remote diabetic retinopathy management using the triad ocular telehealth network," in *The American Telemedicine Association Sixteenth Annual International Meeting and Exposition*, Tampa, Florida, May 2011, p. P. 89.

[14] L. Giancardo, M. D. Abramoff, E. Chaum, T. P. Karnowski, F. Meriaudeau, and K. W. Tobin, "Elliptical local vessel density: A fast and robust quality metric for retinal images," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, Vancouver, Canada, Aug. 2008, pp. 3534 –3537.

[15] T. P. Karnowski, D. Aykac, L. Giancardo, Y. Li, T. Nichols, K. Fox, S. Garg, K. W. Tobin, and E. Chaum, "Combining image and non-image data for automatic detection of retina disease in a telemedicine network," in *Biomedical Sciences and Engineering Conference (BSEC)*, March 2011, pp. 1–4.