

# A Machine Learning Approach for Device-Independent Automated Segmentation of Retinal Cysts in Spectral Domain Optical Coherence Tomography Images

Luis de Sisternes<sup>1</sup>, Jerry Hong<sup>2</sup>, Theodore Leng<sup>3</sup>, Daniel L. Rubin<sup>1,4</sup>

<sup>1</sup>Department of Radiology and <sup>4</sup>Medicine (Biomedical Informatics Research), Stanford University, Stanford, CA, 94305, USA, {ldesist, dlrubin}@stanford.edu  
<sup>2</sup>Palo Alto Senior High School., Palo Alto, CA, 94301, USA, jerryhong1@gmail.com  
<sup>3</sup>Byers Eye Institute at Stanford, Stanford University School of Medicine, Palo Alto, CA, 94303, USA, tedleng@stanford.edu

**Abstract.** We propose a fully automated algorithm to outline cyst locations on 3D spectral domain optical coherence tomography (SD-OCT) data acquired across different vendors. A set of 34 quantitative features are automatically extracted for each voxel in the data, describing its location, intensity, thickness, and histogram statistics within the retinal layer where it is placed, as well as information in neighboring voxels. This set of features is used to train a model that outputs a score representing the probability that the voxel is part of a cyst, using machine learning techniques and a multiple resolution approach. The obtained scores are stratified by an adaptive threshold and further refined by morphological operations to indicate the presence or absence of cyst in each voxel. Evaluation in 15 3D SD-OCT datasets acquired using 4 different vendors in a leave-one-out cross validation scenario, yielded promising results when compared to manual outlines drawn by two readers.

**Keywords.** Retina, SD-OCT, Segmentation, Machine Learning, Retinal Cyst

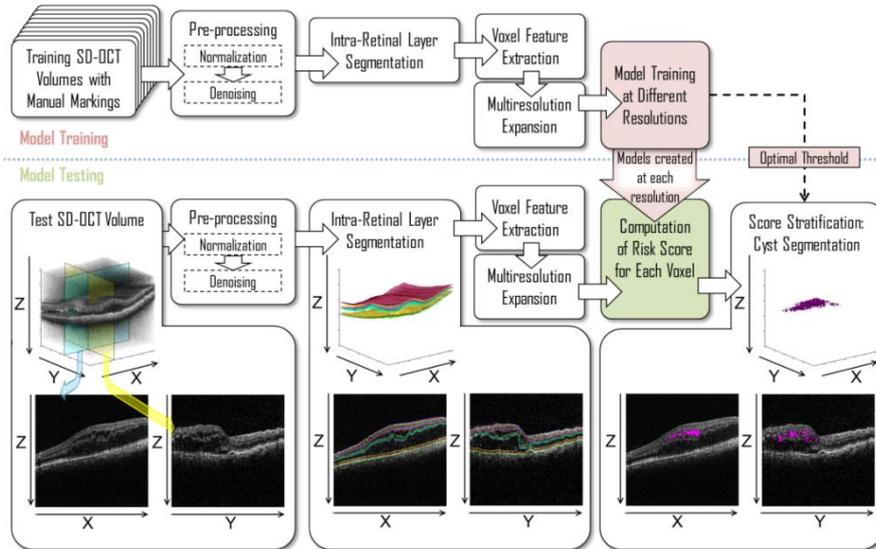
## 1 Introduction

Retinal cysts are a pathological consequence of several common ocular diseases such as retinal diabetic retinopathy, retinal vein occlusion, ocular inflammation and age-related macular degeneration [1,2]. The presence, location, and extent of cystoid macular edema may act as disease biomarkers, thus their volumetric quantification may be beneficial to disease analysis, patient-tailored treatment and treatment progress assessment, in a similar manner that other structures observed in spectral domain optical coherence tomography (SD-OCT) (e.g. drusen) have proven to be useful [3]. Cyst volumes can be manually calculated using SD-OCT [4], but the large amount of imaging data produced in a single SD-OCT volume makes manual assessment extremely laborious, challenging, and prone to differences in subjective interpretation; hence, there is a need for an automated segmentation algorithm to efficiently assess cyst characterization in SD-OCT volumes in an objective manner.

A number of segmentation algorithms have been proposed to date to identify cystoid macular edema and similar closed-contour anomalies like fluid-filled regions and symptomatic exudate-associated derangements [5-12]. However, to our knowledge, none of the previous methods have been developed to work fully automatically and independently of the SD-OCT device used. Most of the previous methods rely on obtaining a coarse result (by either thresholding [6] or using intensity, texture and gradient properties in 2-dimensional B-scan images [9]) to set a baseline for later refinement using level-set approaches and morphological operations [12]. In these approaches, the global parameters, set experimentally, could present limitations when evaluated in images obtained from different vendor devices, where intensity statistics across different intraretinal layers are expected to vary greatly from manufacturer to manufacturer. In order to solve these previous limitations, we describe a fully automated method to segment cystoid macular edema in SD-OCT images across different manufacturers. The novelty of our method lies in employing machine learning techniques in a multiresolution approach, using quantitative imaging features directly extracted from the SD-OCT data. A set of novel features are extracted for each voxel location representing not only information about its intensity, but also its relative position, intensity statistics and axial thickness within each particular intraretinal layer, as well as information in neighboring pixels and location relative to the foveal pit center.

## 2 Methods

A diagram of our automated segmentation approach is displayed in Figure 1, where a description of the coordinates and nomenclature of the SD-OCT data is also shown.



**Fig. 1.** Diagram of steps in the automated cyst segmentation method.

Our method is based in machine learning techniques, where a model is trained using manual markings (to establish the ground truth) and then tested on unseen data. As preprocessing steps, the SD-OCT data are first normalized and denoised using non-local means filtering [13] (radii of the search windows  $t = (14 \mu\text{m}, 82 \mu\text{m})$ ,  $f = (6 \mu\text{m}, 35 \mu\text{m})$  in the axial and horizontal direction, filtering degree  $h = .008$ ). Ten boundaries indicating the axial location of defined intra-retinal layers are then automatically outlined using a developed segmentation algorithm (SOARS: Stanford OCT Automated Retinal Segmentation) [14], from which 7 (describing 6 layers) are considered in the analysis. A set of quantitative features are then extracted to characterize each voxel located between the segmented internal limiting membrane (ILM) and inner segment junction (IS), where we expect the possible cysts to appear. These features describe voxel location within each segmented boundary, intensity value, axial thickness within the retinal layer where it is located, and restricted summed voxel projection (RSVP) [15] statistics within this layer, as well as neighbor information. The set of features are then expanded using a multiresolution approach [16] at 4 possible resolutions to form a set of predictors. These resolution-based predictors are used to train a model that outputs a risk score quantifying the chances of a voxel represents a cyst location using machine learning algorithms [17]. The final segmentation is then obtained by automatically detecting an adaptive threshold to stratify the output scores in those belonging to a cyst or background.

## 2.1 Image dataset

The SD-OCT data used in this work were restricted to the training dataset provided by the OPTIMA laboratory (Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology, Medical University of Vienna) for the Cyst segmentation challenge hosted at MICCAI 2015. We did not consider additional images for training and testing in order to directly compare to other methods being evaluated in the challenge. These data consisted of 15 SD-OCT volumes containing a wide variety of retinal cysts with accompanying clinical ground truth annotation manually drawn by two different readers. The SD-OCT volumes were collected using 4 different vendors at varying resolutions and scanning patterns: Cirrus (Carl Zeiss Meditec, Dublin, CA, USA), Nidek (NIDEK Co., Hiroishi, Gamagori, Japan), Spectralis (Heidelberg Engineering, Heidelberg, Germany) and Topcon (Topcon medical Systems, Santa Clara, CA, USA). Further details about each of the considered volumes characteristics are provided in Table 2 in the results section.

## 2.2 Intra-retinal layer segmentation

We used our previously introduced intra-retinal later segmentation algorithm (SOARS) [14] to automatically outline the axial depth at each location in the X-Y plane of 7 different boundaries located between the ILM and IS: Internal Limiting Membrane, inner and outer boundaries of the Retinal Nerve Fiber Layer, outer boundary of the Inner Plexiform Layer, outer boundary of the Inner Nuclear Layer, outer boundary of the Outer Plexiform Layer and inner boundary of the Inner

Segment / Outer Segment Junction. The location of the center of the foveal pit was also determined automatically, considering the centroid of the expected inner retina thickness depression in a region surrounding the center of the volume.

### 2.3 Feature extraction

A set of 34 quantitative features are extracted for each voxel located within the ILM and the IS. A description of these features is included in Table 1.

**Table 1.** Description of quantitative features extracted for each voxel

Feature Number	Description
1	Recorded voxel intensity normalized across SD-OCT volume
2-7	Indication whether the voxel belongs to the 6 intra-retinal layers
8-14	Axial distance of voxel to the 7 intra-retinal boundaries
15-20	Axial thickness of the 6 intra-retinal boundaries at voxel location
21-26	Recorded intensity in voxels in 3D 6-neighborhood
27	Voxel distance to foveal center
28	Percentile within lower-Gaussian fit of retinal layer intensities
29	Percentile within higher-Gaussian fit of retinal layer intensities
30	Difference significance considering the two Gaussian fits
31	Intensity normalized considering RSVP values
32	Percentile within lower-Gaussian fit of RSVP image intensities
33	Percentile within higher-Gaussian fit of RSVP image intensities
34	Difference significance considering the two RSVP Gaussian fits

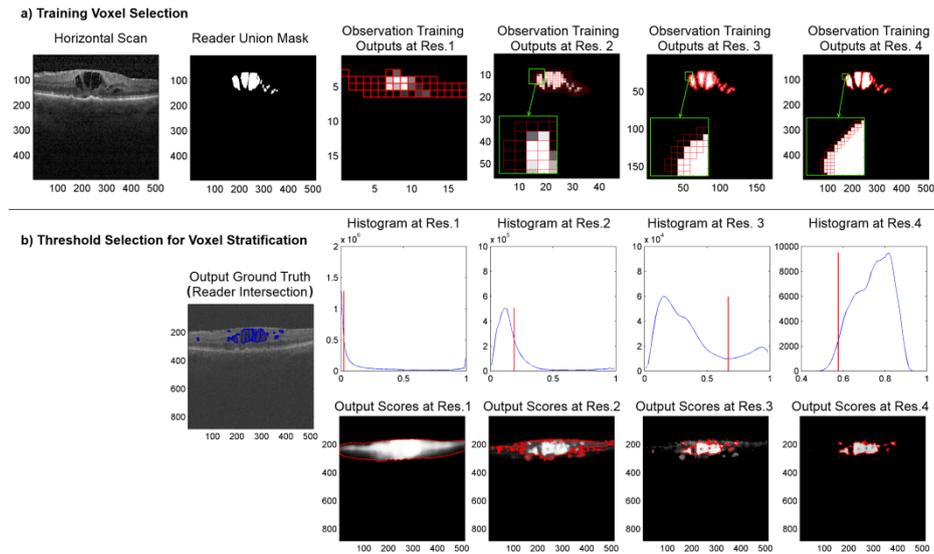
RSVP images were generated projecting the average axial intensity values within the intraretinal layer in which a voxel is located. Features related to lower and higher Gaussian fit of layer intensities (features 28 and 29) or RSVP images (features 32 and 33) refer to percentile of the voxel intensity within the result of fitting two Gaussian distributions to the collection of intensities recorded throughout the intraretinal layer where the voxel is located (for features 28 and 29) or the intensities recorded in the corresponding RSVP image (for features 32 and 33). Difference significance features (features 30 and 34) indicate whether the difference between these two Gaussian distributions is statistically significant and the voxel intensity value is lower than 95% of the higher Gaussian fit values.

### 2.4 Multiresolution predictor expansion

In order to account for the resolution differences across OCT vendors and to also consider features extracted at different resolutions, we resampled the voxel-based extracted features at canonical sizes [16]. We will refer to the resampled voxels as “observations“ and to their resolution-related features as “predictors“. Resampling was done with a Gaussian filter at four canonical possible observation sizes: Res.1 (120,360,360), Res.2 (40,120,120), Res.3 (13.33,40,40) and Res.4 (4.44,13.33,13.33), all in microns in the (X,Y,Z) directions, respectively.

## 2.5 Computation of risk score

We used the predictors extracted at each resolution to create a model generating a score related to the probability that each observation belongs to a cyst. An independent predictive model was created at each of the four resolutions using Lasso regularization [17] in order to prevent overfitting and to enable feature selection. Each model was trained using observations from the training data separated at each observation size within the SD-OCT volume, considering the ratio of voxels outlined as a cyst by any of the two manual readers within the observation span as the ground truth. Due to the large number of observations extracted from each volume, and in an effort to increase training speed, the number of observations at each resolution was also reduced to only consider those not including a ratio of 1 or 0 as ground truth in the immediately higher observation size (lower resolution). This way, each resolution increase can be viewed as a refinement step. This process is indicated in Figure 2a. Output scores were computed for observations at each voxel location of the test data at the four resolutions considered, so as to produce a score for each voxel.



**Fig. 2.** a) Selection of voxels used to train the model at each resolution in and example case (labeled as Spectralis 1 here). The red outlines indicate those voxels that were selected (detail is zoomed in the green squares). b) Threshold selection for stratification at subsequent resolutions in example case (labeled as Topcon 3 here). The red vertical lines in the histograms indicate the selected threshold. The red outlines indicate those voxels with values over the selected threshold at each resolution.

## 2.6 Score stratification and segmentation refinement

The automated process of selecting an SD-OCT volume adaptive threshold for the stratification of output scores is displayed in Figure 2b. From the lowest resolution

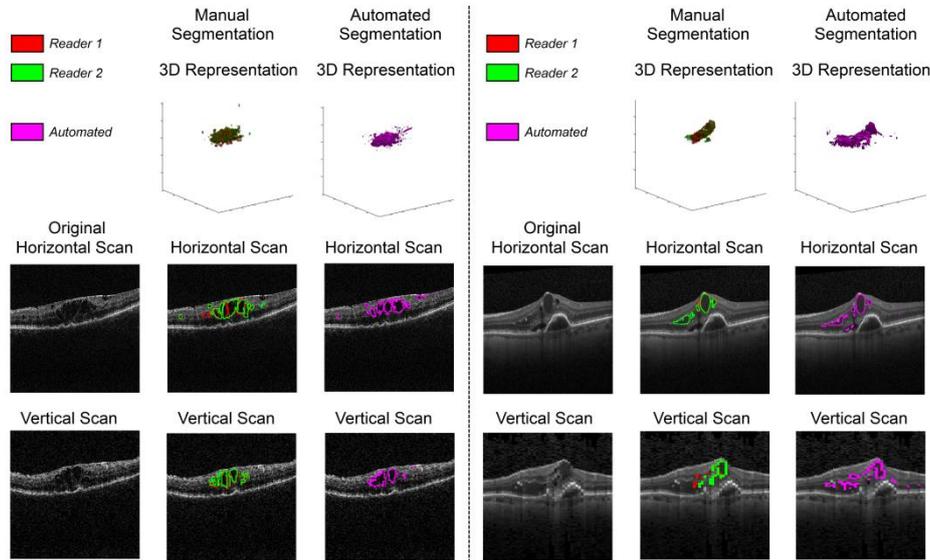
(Res.1) to the highest resolution (Res.4), this process selected as foreground region the output scores over a threshold, and this foreground region was further refined by analyzing the scores at the subsequent higher resolution. Thresholds were selected by analyzing the histogram of scores (computed with 50 bins) at each resolution, limited to the scores selected as foreground in the previous lower resolution. The thresholds for Res.1 and Res.2 were selected as the first histogram location at full-width at half-maximum (FWHM) over the median location of output scores. The thresholds for Res.3 and Res.4 were selected to guarantee the inclusion of 95% of output scores within the higher of two Gaussians fitted to the scores (also limiting that only one Gaussian can be fitted with a mean of less than 0.5). The results from this stratification were later refined by removal of isolated regions (in a 26-neighborhood) of less than 10 voxels, morphological opening by a kernel of radius (45, 15, 15) microns ((X,Y,Z) directions) and filling isolated holes in each B-scan.

### 3 Results

The proposed method was evaluated in the set of provided 15 SD-OCT volumes using a leave-one-out cross-validation scenario. The metric used for evaluation was the Dice coefficient [18], measured between the automated method and each individual manual reader and the intersection of the two readers, computed in four different scenarios: volumetric and B-scan by B-scan values considering the whole SD-OCT volume, and volumetric and B-scan by B-scan values restricted to regions within a central 3mm diameter. We also computed the Dice coefficients between the two manual markings at each scenario for comparison. The results obtained for each of the evaluated volumes and averaged across volumes are summarized in Table 2. Vendor and resolution characteristics for each volume are also indicated. A visualization example for two of the SD-OCT volumes is also displayed in Figure 3.

**Table 2.** Preliminary evaluation results: manual reader comparison (R), comparison between automated method and Reader 1 / Reader 2 / Reader intersection (R1/R2/Inter.).

Volume name	Volume dimensions	Dice coefficient: Volumetric		Dice coefficient: B-scan		Dice coefficient: Volumetric 3mm		Dice coefficient: B-scan 3mm	
		R	R1/R2/Int.	R	R1/R2/Int.	R	R1/R2/Int.	R	R1/R2/Int.
Cirrus 1	1024x512x128	0.69	0.19/0.21/0.19	0.94	0.41/0.40/0.41	0.69	0.27/0.30/0.29	0.87	0.30/0.28/0.30
Cirrus 2	1024x512x128	0.89	0.57/0.57/0.60	0.95	0.47/0.45/0.48	0.89	0.67/0.66/0.71	0.92	0.38/0.35/0.40
Cirrus 3	1024x200x200	0.76	0.57/0.57/0.64	0.68	0.50/0.50/0.52	0.83	0.65/0.65/0.71	0.75	0.48/0.48/0.52
Cirrus 4	1024x512x128	0.65	0.09/0.09/0.09	0.96	0.41/0.41/0.41	0.65	0.30/0.30/0.36	0.91	0.66/0.66/0.67
Nidek 1	512x512x128	0.81	0.77/0.75/0.79	0.84	0.67/0.66/0.69	0.82	0.78/0.76/0.79	0.72	0.49/0.46/0.51
Nidek 2	512x512x128	0.82	0.45/0.44/0.49	0.92	0.64/0.64/0.65	0.84	0.54/0.54/0.58	0.86	0.59/0.58/0.61
Nidek 3	512x512x128	0.66	0.47/0.44/0.49	0.82	0.63/0.62/0.63	0.68	0.62/0.62/0.63	0.84	0.77/0.77/0.77
Spectralis 1	496x512x49	0.80	0.61/0.56/0.62	0.89	0.49/0.47/0.49	0.81	0.63/0.59/0.65	0.81	0.38/0.36/0.40
Spectralis 2	496x512x49	0.90	0.79/0.76/0.81	0.94	0.77/0.75/0.77	0.90	0.81/0.78/0.84	0.89	0.87/0.79/0.87
Spectralis 3	496x1024x49	0.68	0.43/0.42/0.38	0.86	0.50/0.49/0.49	0.69	0.48/0.46/0.42	0.72	0.37/0.35/0.33
Spectralis 4	496x512x49	0.83	0.46/0.45/0.50	0.95	0.79/0.79/0.80	0.84	0.47/0.46/0.51	0.92	0.68/0.67/0.69
Topcon 1	885x512x128	0.68	0.56/0.48/0.51	0.83	0.67/0.64/0.65	0.69	0.68/0.58/0.62	0.83	0.81/0.75/0.78
Topcon 2	885x512x128	0.68	0.41/0.47/0.43	0.75	0.48/0.50/0.49	0.72	0.45/0.52/0.47	0.76	0.48/0.50/0.50
Topcon 3	885x512x128	0.85	0.77/0.78/0.79	0.88	0.62/0.65/0.66	0.86	0.79/0.80/0.81	0.81	0.69/0.73/0.74
Topcon 4	885x512x128	0.81	0.36/0.35/0.40	0.95	0.78/0.78/0.78	0.81	0.36/0.35/0.41	0.90	0.66/0.65/0.67
<b>Average (std)</b>	-	<b>0.77 (0.09)</b>	<b>0.50 (0.20)/ 0.49 (0.19)/ 0.51 (0.21)</b>	<b>0.88 (0.08)</b>	<b>0.59 (0.13)/ 0.58 (0.13)/ 0.59 (0.13)</b>	<b>0.78 (0.08)</b>	<b>0.57 (0.17)/ 0.56 (0.16)/ 0.59 (0.17)</b>	<b>0.83 (0.07)</b>	<b>0.57 (0.18)/ 0.56 (0.17)/ 0.58 (0.18)</b>



**Fig. 3.** Segmentation results for two of the SD-OCT volumes: Left, Nidek 1, shows good quantitative and visual agreement with the manual markings; right, Spectralis 3, did not show good quantitative agreement (Dice measurements) but it shows visually good results.

## 4 Discussion

Our automated cyst segmentation method shows promising results in the evaluated data. The outlines produced by our method had good quantitative resemblance to those produced by the manual readers, although differences were higher than those between the two readers. Upon qualitative review of the automated segmentations, we overserved that they were of very high quality overall, with the exception of one scan where signal quality was poor (Cirrus 1). Taking as example the results displayed on Figure 3 (labeled as Nidek 1 and Spectralis 3 in Table 2), we can observe that while the quantitative comparisons yielded lower values than between the two readers, the outlines marked by our method include regions that could possibly be related to cysts and may have been missed by the manual readers. The ability of the automated method to judge regions in 3D rather than in 2D, B-scan by B-scan (as a manual reader would do), may improve its performance by judging regions that should be included or excluded from the segmentation results. Since there is no absolute gold standard for the segmentations, whether if the method introduced here produces results of higher accuracy than a manual reader is not certain and would require further study.

A limitation of our work is that a small number of images were used for training; better performance could be achieved with a larger set of images. However, obtaining accurate manual markings in a large set of SD-OCT cubes is costly. We also chose not to include additional images in the training set to enable direct comparison of our method to other methods trained and tested using the same set of images.

## References

1. Rotsos, T.G., Moschos, M.M.: Cystoid macular edema. *Clin. Ophthalmol.* 2 (2008) 919-930
2. The Eye Diseases Prevalence Research Group.: The prevalence of diabetic retinopathy among adults in the United States. *Arch. Ophthalmol.* 122 (2004) 552-563
3. Hee, M.R., Puliafito, C.A., Duker, J.S., Reichel, E., Coker, J.G., Wilkins, J.R., Schuman, J.S., Swanson, E.A., Fugimoto, J.G.: Topography of diabetic macular edema with optical coherence tomography. *Ophthalmol.* 105 (1998) 360-370
4. de Sisternes, L., Simon, N., Tibshirani, R., Leng, T., Rubin, D.L.: Quantitative SD-OCT imaging biomarkers as indicators of age related macular degeneration progression. *Invest. Ophthalmol. Vis. Sci.* 55 (2014) 7093-7103
5. Chiu, S.J., Toth, C.A., Rickman, C.B., Izatt, J.A., Farsiu, S.: Automatic segmentation of closed-contour features in ophthalmic images using graph theory and dynamic programming. *Biomed. Opt. Express* 3 (2012) 1127-1140
6. Wilkins, G.R., Houghton, O.M., Oldenburg, A.L.: Automated segmentation of intraretinal cystoid fluid in optical coherence tomography. *IEEE Trans. Biomed. Eng.* 59 (2012) 1109-1114
7. RoyChowdhury, S., Koozekanani, D., Radwan, S., Parhi, K.: Automated Detection and Location Analysis of Diabetic Cysts in Retinal OCT Images: An Iterative Filtering Approach. *Invest. Ophthalmol. Vis. Sci.* 54 (2013) 4901.
8. Wieclawek, W., Pietka, E.: Retina Analysis in Optical Coherence Tomography Images. *Information Technologies in Biomedicine* 3 (2014) 247-258
9. Zhang, L., Zhu, W., Shi, F., Chen, H., Chen, X.: Automated segmentation of intraretinal cystoid macular edema for retinal 3D OCT images with macular hole. *IEEE 12th International Symposium on Biomedical Imaging* (2015) 1494-1497
10. Lang, A., Carass, A., Swingle, E.K., Al-Louzi, O., Bhargava, P., Saidha, S., Ying, H.S., Calabresi, P.A., Prince, J.L.: Automatic segmentation of microcystic macular edema in OCT. *Biomed. Opt. Express* 6 (2014) 155-169
11. Quellec, G., Lee, K., Dolejsi, M., Garvin, M.K., Abramoff, M.D., Sonka, M.: Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in SD-OCT of the macula. *IEEE Trans. Med. Imaging* 29 (2010) 1321-30
12. Zheng, Y., Sahni, J., Campa, C., Stangos, A.N., Raj, A., Harding, S.P.: Computerized assessment of intraretinal and subretinal fluid regions in spectral-domain optical coherence tomography images of the retina. *Am. J. Ophthalmol.* 155 (2013) 277-286
13. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. *Comput. Vis. Patt. Recog.* 2 (2005) 60-65
14. de Sisternes, L., Hu, J., Rubin, D.L., Marmor, M.: Localization of damage in progressive hydroxychloroquine retinopathy on and off the drug: inner versus outer retina, parafovea versus peripheral fovea. *Invest. Ophthalmol. Vis. Sci.* 56 (2015) 3415-3426
15. Rubin, D.L., de Sisternes, L., Kutzscher, L., Chen, Q., Leng, T., Zheng, L.L.: Quantitative evaluation of drusen on photographs. *Ophthalmology.* 120 (2013) 644-644.e2
16. Lindeberg, T.: Scale-space for discrete signals. *IEEE Transactions of Pattern Analysis and Machine Intelligence* (1990) 234-254
17. Tibshirani, R.: Regression shrinkage and selection via the Lasso. *J. Roy. Stat. Soc. B.* 58 (1996) 267-288
18. Sørensen, T.: A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskab* 5 (1948) 1-34.