Published in final edited form as:

Artif Intell Med. 2015 June ; 64(2): 105–115. doi:10.1016/j.artmed.2015.04.002.

Learning from healthy and stable eyes: a new approach for detection of glaucomatous progression

Akram Belghith^a, Christopher Bowd^a, Felipe A. Medeiros^a, Madhusudhanan Balasubramanian^b, Robert N. Weinreb^a, and Linda M. Zangwill^a

Akram Belghith: abelghith@ucsd.edu; Christopher Bowd: cbowd@ucsd.edu; Felipe A. Medeiros: fmedeiros@ucsd.edu; Madhusudhanan Balasubramanian: mbalsbrmn@memphis.edu; Robert N. Weinreb: rweinreb@ucsd.edu; Linda M. Zangwill: lzungwill@ucsd.edu

^aUniversity of California San Diego, Hamilton Glaucoma Center, La Jolla, California, United States 9500 Gilman Drive La Jolla, CA 92093-0946 USA

^bUniversity of Memphis, Department of Electrical and Computer Engineering, Memphis, Tennessee, United States 3815 Central Avenue, Memphis, TN 38152

Abstract

Glaucoma is a chronic neurodegenerative disease characterized by loss of retinal ganglion cells, resulting in distinctive changes in the optic nerve head (ONH) and retinal nerve fiber layer. Important advances in technology for non-invasive imaging of the eye have been made providing quantitative tools to measure structural changes in ONH topography, a crucial step in diagnosing and monitoring glaucoma. Three dimensional (3D) spectral domain optical coherence tomography (SD-OCT), an optical imaging technique, is now the standard of care for diagnosing and monitoring progression of numerous eye diseases.

Method—This paper aims to detect changes in multi-temporal 3D SD-OCT ONH images using a hierarchical fully Bayesian framework and then to differentiate between changes reflecting random variations or true changes due to glaucoma progression. To this end, we propose the use of kernel-based support vector data description (SVDD) classifier. SVDD is a well-known one-class classifier that allows us to map the data into a high-dimensional feature space where a hypersphere encloses most patterns belonging to the target class.

Results—The proposed glaucoma progression detection scheme using the whole 3D SD-OCT images detected glaucoma progression in a significant number of cases showing progression by conventional methods (78%), with high specificity in normal and non-progressing eyes (93% and 94% respectively).

Conclusion—The use of the dependency measurement in the SVDD framework increased the robustness of the proposed change-detection scheme with comparison to the classical support vector machine and SVDD methods. The validation using clinical data of the proposed approach

© 2015 Published by Elsevier B.V.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

has shown that the use of only healthy and non-progressing eyes to train the algorithm led to a high diagnostic accuracy for detecting glaucoma progression compared to other methods.

Keywords

Spatial dependency modeling; copula theory; kernel classifier; glaucoma progression detection

1. Introduction

Glaucoma is an optic neuropathy characterized by distinctive changes in the optic nerve head (ONH) and visual field. Glaucoma is often asymptomatic in its early stages and causes blindness if it remains without treatment. It results in progressive loss of retinal ganglion cells and their axons causing typical changes in the appearance of the retinal nerve fiber layer (RNFL) and the optic disk.

The detection of glaucoma change over time is of high interest in the diagnosis and management of glaucoma particularly for patients as the detection of progression could indicate uncontrolled disease and possible need for therapy advancement. Hence, it is important to develop clinically relevant methods for progression detection in order to avoid permanent damage to the optic nerve head.

In 1851, Helmholtz revolutionized the field of ophthalmology with the invention of the ophthalmoscope, which allowed physicians to examine the disc clinically and identify damages in the optic nerve head associated with glaucoma. This requires clinician identification of the outer and inner borders of the neuroretinal rim and visual estimation of the amount of rim tissue. However, the inner and the outer borders of the ONH neural tissue are not always visible by clinical examination techniques [1]. Furthermore, the clinical examination of the ONH remains subjective, qualitative and with limited reproducibility [2].

In order to assist the expert in qualitative and quantitative analysis, automatic image processing methods have been proposed to facilitate the interpretation of the images obtained by objectively measuring the ONH structure and detecting changes between a reference image (baseline exam) and other images (follow-up exams).

In this context, important advances in technology for non-invasive imaging of the eye have been made providing quantitative tools to measure structural changes in ONH topography, an essential element in detecting glaucoma and monitoring its progression. In particular, the heidelberg retina tomograph (HRT), a confocal scanning laser technology, has been commonly used for glaucoma diagnosis since its commercialization 20 years ago [3]. A limited number of strategies have been proposed that quantitatively and qualitatively detect glaucomatous progression using HRT images. In [4], the topographic change analysis was proposed for assessing glaucomatous changes. This technique has been shown to classify progressing and stable eyes with reasonably good sensitivity and specificity. Another method called the proper orthogonal decomposition [5] indirectly utilizes the spatial relationship among voxels by controlling the family-wise Type I error rate. The Markov random field (MRF) model was used in [6] to model the inter/intra observations dependency allowing a better glaucoma progression detection rate. However, the HRT imaging

technique is limited by its lower resolution (i.e. $300\mu m$ axial resolution for the HRT3). Moreover, because HRT is limited to ONH surface topography, it cannot differentiate between retinal layers. It provides an indirect measure of RNFL thickness that is calculated as the difference between the retinal surface and a standard reference plane $50\mu m$ below the surface of the retina temporal to the ONH.

In contrast, the 3D spectral domain optical coherence tomography (SD-OCT) can differentiate between retinal layers and provide quantitative estimates for change detection. SD-OCT is now the most commonly used instrument for imaging both the ONH and the RNFL thickness. Numerous studies have evaluated glaucoma detection using SD-OCT images. However, most of the studies use the RNFL measurements provided by the commercially available spectral-domain optical coherence tomographers for change detection [7]. Although those methods are successfully applied to SD-OCT images, its use is constrained by specific pre-requisite: it requires an accurate estimation of the RNFL layer thickness. In [8], authors showed that the instrument built-in segmentation software is relatively robust to the image quality and the noise may lower the accuracy of the RNFL layer thickness estimation.

In this paper we propose a hierarchical framework for glaucoma progression detection using 3D Spectralis (Heidelberg engineering) SD-OCT images. This paper is an extended version of the conference paper [9]. Specifically, we explain in more details the change detection algorithm and we add more experiments in the results section. Moreover, we propose the use of a new kernel-based classifier to improve the results of the fuzzy classifier. In contrast to previous works that use the RNFL thickness measurement, we consider the whole 3D volume for progression detection. Our framework is divided into two steps: 1) change detection step which consists of detecting changes between a baseline image and a follow-up image and 2) a classification step which consists of classifying the detected changes into random changes or true changes due to glaucoma progression. For the first step, we propose a fully Bayesian framework for change detection since these methods are relatively simple and offer efficient tools to include *a priori* knowledge through the *a posteriori* probability density function (PDF). In particular, we propose the use of the MRF model to exploit the statistical correlation of intensities among the neighborhood voxels [10]. In order to develop a noise robust algorithm, we propose consideration of the change detection problem as a missing data problem where we jointly estimate the noise hyperparameters and the change detection map. The widely used procedure to estimate the different problem parameters is the Expectation-Maximization (EM) algorithm [11]. However, since we used the MRF model with the change detection map as the prior for the change detection map, the optimization step is intractable. Hence, we propose the use of a Monte Carlo Markov chain (MCMC) technique [12].

Once the change detection map is estimated, we propose the use of kernel-based classifier for glaucoma progression detection. Kernel-based classifiers have several advantages compared with other approaches; they reduce the dimensionality of the data, increase the reliability and the robustness of the method in the presence of noise and allow flexible mappings between objects (inputs) represented by features vectors and class labels (outputs) [13]. As no prior knowledge on glaucoma progression is available, we are interested in the

one class classifier where only the control data (healthy eyes and stable glaucoma eyes) are used to train the classifier. To this end, we have elected to use the support vector data description (SVDD) method [14] here. The SVDD one class-classifier method maps the data into a high-dimensional feature space. In this new space, a hypersphere that encloses most of the dataset belonging to the class of interest (the *target* class). Although basic kernel functions can be successfully applied for change detection [15–17], they do not exploit the additional constraints that are often available, such as the dependencies and the distribution of different features. We show in this paper that the classification should be more efficient if such information is integrated. To account for these characteristics in our change-detection scheme, we propose the use of a new kernel function that combines some properties of the old kernel functions with new information about the feature distribution and dependencies [17].

The paper is divided into three methodology sections. In section 2, the proposed change detection scheme is presented. In section 3, we describe the classification step. Then, in section 4 results obtained by applying the proposed scheme to clinical data is presented. The diagnostic accuracy of this novel proposed approach is compared to two existing progression detection RNFL based approaches: the artificial neural network classifier (ANN) and the support vectors machine (SVM) classifier.

2. Change detection

2.1. Direct model

Let us consider the detection of changes in a pair of amplitude images. We denote by $I_0 = \{I_0(i) | i = 1, \dots, M\}$ and $I_1 = \{I_1(i) | i = 1, \dots, M\}$, where M is the number of the image voxels, two images acquired over the same eye at times t_0 and t_1 , respectively ($t_1 > t_0$), and coregistered. In this work, we assume that the noise is additive, white and normally distributed.

The Spectralis SD-OCT instrument features three different options to enhance reproducibility and reduce the noise. A real time eye-tracking device (eye tracker) compensates for involuntary eye movements during the scanning process, a retest function assures that follow-up measurements are taken from the same area of the retina as the baseline examination and a Heidelberg noise reduction option that average automatically several images (10 images are the number recommended by the manufacturer) at the same location to increase image signal to noise ratio (SNR) and improve the quality of subsequent images. As recommended by the manufacturer, only images with a signal quality of equal or greater than 20 dB were used in this study. The mean and the standard deviation of different image qualities are 26 dB and 3 respectively (range 20–40 dB). In this study, five images have a SNR equals to 40 dB (1% of the whole number of images). We assume in these images that most of the noise was removed by the instrument. These images are then used to study the distribution of the retinal tissue intensities. Four distribution candidates: 1) gamma, 2) log-normal, 3) Weibull and 4) normal distributions were used to fit the retinal tissue intensities. Figure 1, presents the fitting results of the different distributions. Using the Bayesian information criterion (BIC) [18], the gamma distribution was identified as the best candidate. It is important to note that the BIC criterion does not allow to test if one model is

statistically better than the other one. In order to overcome this problem, the F-test can be used [18]. However, this test can only compare two nested models (i.e; one model is a special case of the other one). To this end, we used the generalized gamma distribution to test if adding a new parameter will significantly increase the goodness of fit. Note that the Weibull distribution and the log-normal distribution are also special cases of the generalized gamma distribution. Our simulations showed that there is no statistical difference between gamma and generalized gamma distributions. Therefore, the gamma distribution with the hyperparameters ($\alpha > 0, \beta > 0$), denoted by $\mathcal{G}(\alpha, \beta)$, is used in this paper to fit the retinal tissue intensities. The direct model for both images I_0 and I_1 is then given by:

$$I_0 = X_0 + N_0; \quad I_1 = X_1 + N_1 \quad (1)$$

where $X_0 = \{X_0(i) | i = 1, \dots, M\}$, $X_0(i) \sim \mathcal{G}(\alpha_0, \beta_0)$ and $X_1 = \{X_1(i) | i = 1, \dots, M\}$, $X_1(i) \sim \mathcal{G}(\alpha_1, \beta_1)$ are the noise free 3D SD-OCT ONH images and $N_0 = \{N_0(i) | i = 1, \dots, M\}$, $N_0(i) \sim \mathcal{N}(0, \sigma_0)$ and $N_1 = \{N_1(i) | i = 1, \dots, M\}$, $N_1(i) \sim \mathcal{N}(0, \sigma_1)$ are additive, white and normally distributed noises. The change-detection problem is formulated as a hypothesis testing problem by denoting the "no-change", the "increase change" (i.e; ONH hypertrophy) and the "decrease change" (i.e; ONH atrophy) hypotheses as H_0 , H_1 and H_2 respectively.

Two widely used approaches for the change detection can be distinguished: the image-difference approach and the image-ratio approach. On the one hand, the image-difference approach leads to the following direct model: $I_0 - I_1 = (X_0 - X_1) - (N_0 - N_1)$. The advantage of the image difference approach is that the resulting noise is normally distributed, as the difference of two normal distribution is a normal distribution as well, which allows an easy modeling of the likelihood. However, a given site of the free noise image difference at the position $i \in \{1, \dots, M\}$ ($i = (X_0(i) - X_1(i))$) follows a gamma difference distribution:

$$p(\Delta(i); \alpha_0, \beta_0, \alpha_1, \beta_1) = \int_0^\infty \mathcal{G}(X_0(i), \alpha_0, \beta_0) \mathcal{G}(X_0(i) - \Delta(i), \alpha_1, \beta_1) dX_0(i) \quad (2)$$

In case where $\alpha_0 = \alpha_1 = \alpha$, $p(\Delta(i); \alpha, \beta_0, \beta_1)$ is given by:

$$p(\Delta(i); \alpha, \beta_0, \beta_1) = \frac{(1-c^2)^{a+\frac{1}{2}} (|\Delta(i)|)^a}{\sqrt{\pi} \Gamma(a+\frac{1}{2}) 2^a b^{a+1}} \exp\left(-\frac{c\delta(i)}{b}\right) B\left(\frac{-\delta(i)}{b}; a\right) \quad (3)$$

where $B(., a)$ is the modified Bessel function of the second kind and of order a ,

$$\Gamma(y) = \int_0^\infty t^{y-1} \exp^{-t} dt, \quad a = \alpha - \frac{1}{2}, \quad b = 2 \frac{\beta_1 \beta_2}{\beta_1 + \beta_2} \quad \text{and} \quad c = \frac{\beta_2 - \beta_1}{\beta_1 + \beta_2}.$$

Proof: Let x_0 and x_1 be two independent random variables following a gamma distribution $\mathcal{G}(\alpha_0, \beta_0)$ and $\mathcal{G}(\alpha_1, \beta_1)$:

$$p(x_0, \alpha_0, \beta_0) = \frac{x_0^{\alpha_0-1}}{\Gamma(\alpha_0) \beta_0^{\alpha_0}} \exp\left(-\frac{x_0}{\beta_0}\right), \quad p(x_1, \alpha_1, \beta_1) = \frac{x_1^{\alpha_1-1}}{\Gamma(\alpha_1) \beta_1^{\alpha_1}} \exp\left(-\frac{x_1}{\beta_1}\right) \quad x_0, x_1 > 0 \quad (4)$$

The PDF of the difference $\delta = (x_0 - x_1)$ is given by:

$$\begin{aligned} p(\delta; \alpha_0, \beta_0, \alpha_1, \beta_1) &= \int_0^\infty \mathcal{G}(x_0, \alpha_0, \beta_0) \mathcal{G}(x_0 - \delta, \alpha_1, \beta_1) dx_0 \\ &= \frac{1}{\Gamma(\alpha_0)\beta_0^{\alpha_0}\Gamma(\alpha_1)\beta_1^{\alpha_1}} \int_0^\infty x_0^{\alpha_0-1} \exp\left(-\frac{x_0}{\beta_0}\right) (x_0 - \delta)^{\alpha_1-1} \exp\left(-\frac{(x_0 - \delta)}{\beta_1}\right) dx_0 \end{aligned}$$

In case where $\alpha_0 = \alpha_1 = \alpha$, $p(\delta; \alpha, \beta_0, \beta_1)$ is given by:

$$p(\delta; \alpha, \beta_0, \beta_1) = \frac{\exp\left(-\frac{\delta}{\beta_1}\right)}{\Gamma(\alpha)^2(\beta_0\beta_1)^\alpha} \int_0^\infty x_0^{\alpha-1} (x_0 - \delta)^{\alpha-1} \exp(-\varphi x_0) dx_0 \quad (5)$$

where $\varphi = \frac{1}{\beta_0} + \frac{1}{\beta_1}$. Or from [19], we have

$$\int_0^\infty x_0^{\alpha-1} (x_0 - \delta)^{\alpha-1} \exp(-\varphi x_0) dx_0 = \frac{1}{\sqrt{\pi}} \left(\frac{|\delta|}{\varphi}\right)^{\alpha-\frac{1}{2}} \exp\left(-\frac{\delta\varphi}{2}\right) \Gamma(\alpha) B\left(-\frac{\delta\varphi}{2}; \alpha - \frac{1}{2}\right) \quad (6)$$

where $B(\cdot; \alpha - \frac{1}{2})$ is the modified Bessel function of the second kind and of order $\alpha - \frac{1}{2}$. Let us define a , b and c as: $a = \alpha - \frac{1}{2}$, $b = 2\frac{\beta_1\beta_2}{\beta_1 + \beta_2}$ and $c = \frac{\beta_2 - \beta_1}{\beta_1 + \beta_2}$. The gamma difference distribution is then given by:

$$p(\delta; \alpha, \beta_0, \beta_1) = \frac{(1-c^2)^{a+\frac{1}{2}} (|\delta|)^a}{\sqrt{\pi} \Gamma(a+\frac{1}{2}) 2^a b^{a+1} \exp(-\frac{c\delta}{b})} B\left(\frac{-\delta}{b}; a\right) \quad (7)$$

Hence, the estimation of gamma difference distribution hyperparameters is complicated and therefore the image difference approach is not well suited to our problem. On the other hand, the image-ratio approach leads to the following direct model: $R = \frac{I_0}{I_1} = \frac{X_0 + N_0}{X_1 + N_1}$, $R = \{R(i) | i = 1, \dots, M\}$. This model is unfortunately intractable. To overcome this problem, we propose a hierarchical change detection framework which consists of estimating the noise free SD-OCT images $\hat{X}_0$ and $\hat{X}_1$ and then use the image ratio approach for change detection. The direct model is then given by $R = \frac{\hat{X}_0}{\hat{X}_1}$. The gamma ratio distribution is expressed as:

$$p(R(i); \alpha_0, \beta_0, \alpha_1, \beta_1) = \frac{\Gamma(\alpha_0 + \alpha_1)}{\Gamma(\alpha_0)\Gamma(\alpha_1)} R(i)^{\alpha_0-1} \left(\frac{\beta_0}{\beta_1}\right)^{\alpha_0} \left(R(i) + \frac{\beta_0}{\beta_1}\right)^{-(\alpha_0 + \alpha_1)} \quad (8)$$

Proof: Let x_0 and x_1 be two independent random variables following a gamma distribution $\mathcal{G}(\alpha_0, \beta_0)$ and $\mathcal{G}(\alpha_1, \beta_1)$. The PDF of the ratio $r = (x_0/x_1)$ is given by:

$$\begin{aligned}
p(r; \alpha_0, \beta_0, \alpha_1, \beta_1) &= \int_0^\infty \mathcal{G}(x_1 r, \alpha_0, \beta_0) \mathcal{G}(x_1, \alpha_1, \beta_1) x_1 dx_1 \\
&= \frac{r^{\alpha_0-1}}{\Gamma(\alpha_1)\beta_1^{\alpha_1}} \frac{1}{\Gamma(\alpha_0)\beta_0^{\alpha_0}} \left(\frac{r}{\beta_0} + \frac{1}{\beta_1}\right)^{\alpha_0+\alpha_1} \Gamma(\alpha_1+\alpha_0) \times \int_0^\infty \frac{1}{\left(\frac{r}{\beta_0} + \frac{1}{\beta_1}\right)^{\alpha_0+\alpha_1} \Gamma(\alpha_1+\alpha_0)} x_1^{\alpha_0+\alpha_1-1} \exp\left(-x_1 \left(\frac{r}{\beta_0} + \frac{1}{\beta_1}\right)\right) dx_1 \quad (9) \\
&= \frac{\Gamma(\alpha_0+\alpha_1)}{\Gamma(\alpha_0)\Gamma(\alpha_1)} r^{\alpha_0-1} \left(\frac{\beta_0}{\beta_1}\right)^{\alpha_0} \left(r + \frac{\beta_0}{\beta_1}\right)^{-(\alpha_0+\alpha_1)}
\end{aligned}$$

since the PDF integrates to 1.

In this work, we assume that $\alpha_0 = \alpha_1 = \alpha$ and $\beta = \frac{\beta_0}{\beta_1}$, $p(R(i); \alpha, \beta)$ is given by;

$$p(R(i); \alpha, \beta) = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} R(i)^{\alpha-1} \beta^\alpha (R(i) + \beta)^{-2\alpha} \quad (10)$$

Indeed, the scale parameters α_0 and α_1 measure how spread out the gamma distribution is. The larger the scale parameters, the more the spread out of the gamma distribution. Because glaucoma progression is often slow, the observed changes between two visits are relatively small. α_0 was taken equal to α_1 for sake of model simplicity. It is easy to extend the change detection scheme to the case of $\alpha_0 \neq \alpha_1$, particularly when we have long follow-up duration, by considering the non simplified gamma ratio distribution as follows:

$$p(R(i); \alpha_0, \alpha_1, \beta) = \frac{\Gamma(\alpha_0+\alpha_1)}{\Gamma(\alpha_0)\Gamma(\alpha_1)} R(i)^{\alpha_0-1} (\beta)^{\alpha_0} (R(i) + \beta)^{-(\alpha_0+\alpha_1)} \quad (11)$$

In contrast to the gamma difference distribution, the estimation of the simplified gamma ratio distribution hyperparameters is easy to perform. Finally, the change detection is handled through the introduction of change class assignments Q . To introduce a spatial *a priori* knowledge on $Q = \{Q(i) | i = 1, \dots, M\}$, we used the MRF model so that the change status $Q(i)$ of a voxel $R(i)$ depends on the change status of its neighborhood.

2.2. The estimation scheme

The Bayesian model aims at estimating the model parameters (X_0, X_1, Q) and hyperparameters Θ . This requires the definition of the likelihood and prior distributions. We now present each term of the hierarchical Bayesian model.

Likelihood—The definition of the likelihood depends on the noise model. As we assumed that the noise in both images I_0 and I_1 is white and normally distributed with standard deviations σ_0 and σ_1 respectively, the likelihood is then given by:

$$p(I_0 | X_0; \sigma_0) = \mathcal{N}(0, \sigma_0); p(I_1 | X_1; \sigma_1) = \mathcal{N}(0, \sigma_1) \quad (12)$$

Model prior—The prior on X_0 and X_1 is given by the gamma density $\mathcal{G}$:

$$p(X_0; \alpha_0, \beta_0) = \prod_{i=1}^M \mathcal{G}(x_0(i), \alpha_0, \beta_0); \quad p(X_1; \alpha_1, \beta_1) = \prod_{i=1}^M \mathcal{G}(x_1(i), \alpha_1, \beta_1) \quad (13)$$

To describe the marginal distribution of R conditioned to $H_{l,l \in \{0,1,2\}}$, we used the ratio gamma distribution:

$$p(R|Q=H_l; \theta_l, \kappa_l) = \prod_{i=1:M} \frac{\Gamma(2\theta_l)}{\Gamma(\theta_l)^2} r(i)^{\theta_l-1} \kappa_l^{\theta_l} (r(i) + \kappa_l)^{-2\theta_l} \quad (14)$$

Concerning the change variable, we assume that $p(Q; \Theta)$ is a spatial MRF prior. In this study we adopt the Potts model with interaction parameter ζ :

$$p(Q; \zeta) = \frac{1}{Z} \exp(-\zeta U(Q)), \quad U(Q) = \sum_{i \sim j} \delta(q_i, q_j) \quad (15)$$

where Z is the normalization constant and d the delta Kroneker function. Note that we have opted for the 6-connexity 3D neighboring system. The whole set of the model hyperparameters is then given by $\Theta = \{\sigma_{j=0,1}, \alpha_{j=0,1}, \beta_{j=0,1}, \theta_{l=0:2}, \kappa_{l=0:2}, \zeta\}$. Because no knowledge on the noise variance is available, we retained the following prior distributions

for σ_0 and σ_1 : $p(\sigma_0^2) = \frac{1}{\sigma_0}$, $p(\sigma_1^2) = \frac{1}{\sigma_1}$. The non-negativity of the hyperparameters $\{\alpha_{j=0,1}, \beta_{j=0,1}, \kappa_{l=0:2}, \theta_{l=0:2}, \zeta\}$ is guaranteed through the use of the exponential densities:

$$p(\alpha_0, \beta_0) = \frac{1}{\eta} \exp\left(-\frac{\alpha_0}{\eta}\right) \frac{1}{\vartheta} \exp\left(-\frac{\beta_0}{\vartheta}\right), \quad p(\alpha_1, \beta_1) = \frac{1}{\eta} \exp\left(-\frac{\alpha_1}{\eta}\right) \frac{1}{\vartheta} \exp\left(-\frac{\beta_1}{\vartheta}\right), \quad p(\theta_l, \kappa_l) = \frac{1}{\eta} \exp\left(-\frac{\theta_l}{\eta}\right) \frac{1}{\varsigma} \exp\left(-\frac{\kappa_l}{\varsigma}\right)$$

and $p(\zeta) = \frac{1}{\kappa} \exp\left(-\frac{\zeta}{\kappa}\right)$. Note that the values of $\{\eta, \vartheta, \varsigma, \kappa\}$ are fixed empirically. The choice of these values do not influence the results but has an impact on the speed of the algorithm convergence [20]. By combining the likelihood and the prior knowledge using the **Bayes' rule** and giving our hierarchical approach we adopted, the obtained posterior distribution $p(Q, \Theta|R)$ is intractable for our model. Hence, we propose the use of a MCMC procedure to estimate the model parameters and hyperparameters. The principle of MCMC method is to generate samples drawn from the posterior densities and then to be able to achieve parameter estimation. We use a Gibbs sampler based on a stationary ergodic Markov chain allowing to draw samples whose distribution asymptotically follows the *a posteriori* densities.

2.3. Inference scheme

The Gibbs sampler decomposes the problem of generating a sample from a high dimension PDF by simulating each variable separately according to its conditional PDF. Since no classical expressions for the posterior PDFs are available, there is no direct simulation algorithm for them. To overcome this problem, we perform Hastings-Metropolis steps within the Gibbs sampler [12] to ensure that all convergence properties are preserved [21]. Moreover, since no knowledge on the posterior definition is available, we have opted for the random walk version of the Hastings-Metropolis algorithm. The proposal distribution p is

then defined as $p(.|Y) = Y + p'(.)$ where p' is a distribution which does not depend on Y and is usually centered on 0. Hence, at each iteration, a random move is proposed from the actual position. The choice of a good proposal distribution is crucial to obtain an efficient algorithm, and the literature suggests that, for low dimension variables, a good proposal should lead to an acceptance rate of 0.5 [12]. The determination of a good proposal distribution can be solved by using a standard Gaussian distribution with an adaptive scale technique [22]. The main Gibbs sampler steps are described in Algorithm. 1. We used a burn-in period of $h_{min} = 500$ iterations followed by another 1000 iterations for convergence ($h_{max}=1500$). The change detection map Q is estimated using the maximum a posteriori

MAP estimator: $\hat{Q} = \arg\max_{H_i} \bar{p}_{H_i}$, where

$$\bar{p}_{H_1} = \frac{1}{h_{max} - h_{min}} \sum_{h=h_{min}+1}^{h_{max}} p^{[h]}(Q=H_1 | R^{[h]}, \Theta^{[h-1]}), \bar{p}_{H_2} = \frac{1}{h_{max} - h_{min}} \sum_{h=h_{min}+1}^{h_{max}} p^{[h]}(Q=H_2 | R^{[h]}, \Theta^{[h-1]})$$

and $p_{H_3} = 1 - p_{H_1} - p_{H_2}$.

Algorithm 1

Sampling algorithm

-
- 1 Initialization of $X_0^{[0]}, X_1^{[0]}, Q^{[0]}$ and $\Theta^{[0]}$
 - 2 For each iteration h repeat:
 - i. Sampling $X_0^{[h]}$ from $p(X_0 | I_0, \Theta^{[h-1]})$
 - ii. Sampling $X_1^{[h]}$ from $p(X_1 | I_1, \Theta^{[h-1]})$
 - iii. Calculating $R^{[h]} = \frac{X_0^{[h]}}{X_1^{[h]}}$
 - iv. Creating a configuration of Q basing on $R^{[h]}$
 - v. Calculating $p^{[h]}(Q=H_i | R^{[h]}, \Theta^{[h-1]})$
 - vi. Sampling $\Theta^{[h]}$ from $p(\Theta | X_0^{[h]}, X_1^{[h]}, I_0, I_1)$
 - vii. Until Convergence criterion is satisfied
-

3. Classification

This step aims at classifying an image into non-progressing or the progressing classes based on the estimated change detection map. As no prior knowledge on glaucoma progression is available, we are interested in the one class classifier where only the control data (healthy eyes and stable glaucoma eyes) are used to train the classifier. To this end, we have elected to use the SVDD method [14] here. The SVDD enables us to distinguish between targets and outliers by defining a closed boundary around the target data. This is equivalent to map a data set defined over the input space into a higher-dimensional Hilbert space (the feature space) then to draw a minimum-volume hypersphere in the kernel feature space that includes all or most of the target objects.

Let $\{\mathbf{x}_i\}_{i=1 \dots K}$, with $\mathbf{x}_i \in \mathbb{R}^N$, be the vector containing the N features of a given object (in our case the estimated change detection map). Kernel methods compute the similarity between samples $\mathbf{x}_i$ using pairwise inner products taken between mapped samples. The most common

kernels functions are the linear kernel $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \cdot \mathbf{x}_j$ and the radial basis function (RBF) $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2)$. These kernel functions do not exploit additional constraints such as the feature dependency (a measurement that models the *distance to the independence* for two random variables) and thus make the change detection less robust. Consider two signals that are both realizations of random processes. If they are affected by the same linear change, both random fields will tend to increase or decrease with the same amplitude. In this work, we opted for Gaussian copula to model non-linear dependency for the multivariate case [17] which has been shown to be effective for treating dependency [23]. In order to include the *distance between features* within the new kernel function, we use the RBF function, which offers some degrees of freedom because of the standard deviation hyperparameter σ with information about the distance to the independence of features using the copula theory [17]. Thus, we adopt the Gaussian copula, which is given here: $\forall \mathbf{y} = (y_1, \dots, y_L) \in \mathbb{R}^L$,

$$c_G(\mathbf{y}) = |\Gamma|^{-\frac{1}{2}} \exp \left[-\frac{\tilde{\mathbf{y}}^T (\Gamma^{-1} - I) \tilde{\mathbf{y}}}{2} \right] \quad (16)$$

where $\tilde{\mathbf{y}} = (\Phi^{-1}(y_1), \dots, \Phi^{-1}(y_L))^T$, in which $\Phi(\cdot)$ the standard Gaussian cumulative distribution function, Γ is the inter-data correlation matrix, and I is the $L \times L$ identity matrix. The proposed kernel function is

$$\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = (E[C_G(\mathbf{x}_i, \mathbf{x}_j)]) \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2) \quad (17)$$

where $C_G(\mathbf{x}_i, \mathbf{x}_j) = [c_G(x_i(1), x_j(1)), \dots, c_G(x_i(K), x_j(K))]$, in which K is the length of the vector $\mathbf{x}_i$. As the dependency of a couple $(\mathbf{x}_i, \mathbf{x}_j)$ increases, the $E[C_G(\mathbf{x}_i, \mathbf{x}_j)]$ value approaches 1. The inter-data correlation matrix Γ is estimated using the maximum-likelihood estimator [24]. Note that the new kernel function satisfies Mercer's condition [17].

Once we established the kernel function, we turn now to the classification step which consists of estimating the optimal hyper-sphere that includes all or most of the target objects. The sphere is characterized by its center $\mathbf{a}$ and its radius $R > 0$. Minimizing the volume of the sphere reduces to minimizing R^2 with the constraints [14]: $\kappa(\mathbf{x}_i - \mathbf{a}, \mathbf{x}_i - \mathbf{a}) \leq R^2 \forall i$. To allow for the possibility of outliers in the target set, the distance from $\mathbf{x}_i$ to the center $\mathbf{a}$ should strictly not be smaller than R^2 and larger distances should be penalized. Therefore, we introduce slack variables $\zeta_i \geq 0$, and the minimization problem becomes

$\min_{R, \mathbf{a}, \zeta_i} \{R^2 + C \sum_i \zeta_i\}$ with the constraint that almost all objects that belong to the target class should be within the sphere: $\kappa(\mathbf{x}_i - \mathbf{a}, \mathbf{x}_i - \mathbf{a}) \leq R^2 + \zeta_i \forall i, \zeta_i \geq 0$. The parameter C controls the trade-off between the volume and the errors. The leave-one-out cross-validation estimator was used to estimate our model hyperparameters $(R, \mathbf{a}, \zeta_i)$ [25]. This algorithm, often cited as being highly attractive for the purposes of model selection, provides an almost unbiased estimate.

The selection of the features depends on the eye disease. The typical markers of the glaucoma progression are the RNFL thinning [26] and the rim area decrease [27]. Therefore, we focused on these areas to classify an eye into progressing and non-progressing (stable) classes. Figure 2 shows the RNFL and the rim area in a B-scan. The estimation of this area for each B-scan is addressed using the method presented in [28]. As in [5, 9], we considered two features as input for the classifier: 1) *feature1*: the number of changed RNFL and rim area locations and 2) *feature2*: the ratio image intensity R of changed RNFL and rim area locations. In this work, only the loss of retinal height in neighboring areas is considered change due to glaucomatous progression because an increase in retinal height is considered improvement (possibly due to treatment or tissue remodeling). The benefits of the SVDD classifier lies in its ability to only use the target class (non-progressing class) for training in contrast to the SVM, for example, where the information about the outlier class (progressing) are needed in the training step. Figure 3 presents an example of a hypersphere enclosing most of the targets estimated using the SVDD approach and figure 4 presents an example of a hyperplane separating most of the targets from the outliers estimated using the SVM approach.

4. Experiments

This section aims at evaluating the proposed framework. The different independent datasets used for validation are described in the sub-section 4.1. Change detection results on simulated and semi-simulated datasets are presented in sub-section 4.2. Finally, the classification results on clinical datasets are presented in sub-section 4.3.

4.1. Datasets for experiments

The proposed framework was experimentally validated on simulated, semi-simulated and clinical datasets. Our clinical datasets consists of 117 eyes from 75 participants. Only patients with primary open angle glaucoma were selected. Specifically, the first clinical dataset consists of 27 eyes from 27 participants with glaucoma progression. The glaucomatous progression was defined based on 1) likely progression by visual field (based on longitudinal standard automated perimetry testing) or 2) progression by stereo-photographic assessment of the optic disk. Progressive changes in the stereo-photographic appearance of the optic disk between baseline and the most recent available stereo-photograph (patient name and diagnosis were masked) were assessed by two observers based on a decrease in the neuroretinal rim width, appearance of a new RNFL defect, or an increase in the size of a pre-existing RNFL defect. A third observer adjudicated any differences in assessment between these two observers. For the selected 27 patients, the two observers agreed on 26 eyes (96 %) as progressing glaucoma and the third observer assessed the remaining eye (4 %). The reason for this is no ground truth for glaucomatous change in the 3D SD-OCT images is available. The second clinical dataset consists of 40 eyes from 23 healthy participants with no history of IOP > 22 mmHg, normal appearing optic disk by stereo-photography and visual field exams within normal limits. The third clinical dataset consists of 50 eyes from 26 participants considered stable (each participant was imaged once a week for five consecutive weeks). All eligible participants were recruited from the university of California, San Diego diagnostic innovations in glaucoma study (DIGS) with at

least 5 good quality visual field exams to ensure an accurate diagnosis. Note that for each exam, the baseline image and the follow-up image are co-registered using built in instrument software and saved together. Table 1 summarizes the clinical dataset.

Because no ground truth for glaucomatous change in the 3D SD-OCT images is available, we generated a simulated and a semi-simulated datasets to assess the proposed change detection method. The simulated dataset consists of one hundred images with different values of PSNR (30, 25, 20 and 15 dB) ($PSNR = 10 \log_{10}(\max(Image)^2/E[(Noise)^2])$). Each simulated image is the ratio of two synthetic OCT images. Fig 5 presents an example of a simulated image. The semi-simulated dataset was constructed from one hundred normal 3D SD-OCT images. Changes were simulated 1) by the permuting 1%, 2%, 3% and 5% of image regions in the images, respectively 2) by modifying the intensities of each $15 \times 15 \times 3$ voxel-sized regions in the images randomly with 0.5%, 0.75%, 1% and 1.25% probabilities, respectively. The intensities were modified by randomly multiplying the real intensities by 0.5, 0.75, 1.5 or 2 with a 0.25% probability. In order to emphasize the robustness of the proposed approach, an additive Gaussian noise with different values of peak signal to noise ratio (PSNR) (30, 25, 20 and 15 dB) was added to both the original image and the semi-simulated images. Fig 6 presents an example of a semi-simulated image. Table 2 summarizes the simulated and semi-simulated datasets.

4.2. Change detection algorithm assessment

To the best of our knowledge, this is the first algorithm to consider the whole 3D SD-OCT images for glaucoma change detection. Therefore, there are no glaucoma change detection algorithms that can be used to evaluate the proposed algorithm. To this end, and in order to evaluate the benefit of the proposed change detection algorithm and particularly the use of the MRF model to handle voxel spatial dependency, we consider two methods that were primarily used to detect changes in the remote sensing synthetic-aperture radar (SAR) images. The intuition behind this choice is that we assume in both images that the intensities of pixels follow a gamma distribution in particular for low-resolution SAR images [29]. Among the different methods for change detection in SAR images, we selected:

- The kernel k-means algorithm [30] with the RBF Gaussian kernel,
- A Bayesian threshold-based method [31],

These methods are robust against the noise and have a good change detection rate compared with existing methods. To perform the change detection evaluation, we use the Percentage of False Alarm PFA , the Percentage of Missed Detection PMD and the Percentage of Total Error PTE measurements defined by:

$$PFA = \frac{FA}{N_F} \times 100\%; \quad PMD = \frac{MD}{N_M} \times 100\%; \quad PTE = \frac{MD + FA}{N_M + N_F} \times 100\%$$

where FA stands for the number of unchanged voxels that were incorrectly determined as changed ones, N_F the total number of unchanged voxels, MD the number of changed voxels that were mistakenly detected as unchanged ones, and N_M the total number of changed voxels. The kappa coefficient of agreement between the ground truth and the change

detection map was also calculated. The kappa statistic was originally developed by Cohen [32] in order to take into account the amount of agreement that could be expected to occur through chance. A detailed description of the kappa coefficient of agreement can be found [33]. The paired t-test was used to assess the significance of the change detection result difference. The Table 3 shows the results of the change detection on simulated and semi-simulated datasets. The Bayesian Markovian approach ($\kappa=0.78$) performs statistically better than the kernel k-means ($\kappa=0.71$) and the Bayesian threshold-based ($\kappa=0.64$) methods. This means that the proposed MRF *priori* we considered improves the change detection results by taking the spatial dependency of voxels into the detection scheme. In order to evaluate the benefit of the MCMC used in the change detection scheme, we compared the proposed estimation procedure with two other procedures:

- The stochastic expectation-maximization (SEM) procedure [34],
- The iterative conditional mode (ICM) [35],

The change detection results obtained on the simulated dataset in Table 4. As can be seen, the proposed MCMC procedure tends to perform better than the SEM and the ICM procedures. This means that our estimation procedure is robust to the noise. Note the results on the semi-simulated datasets are similar to those obtained on the simulated dataset, so the semi-simulated results are omitted for the sake of brevity.

4.3. Classification results

This section describes the glaucoma progression detection results obtained with the proposed scheme. In order to compare the diagnostic accuracy of the classification methods, we calculated the area under the receiver operating characteristic AUROC, the *sensitivity* and the *specificity* measurements [28]. The proposed framework was experimentally validated with clinical datasets. In order to evaluate the benefit of the proposed glaucoma progression detection scheme, we have compared the proposed framework called as **Bayesian-kernel detection scheme (BKDS)** with six other methods:

1. The SVM classifier of the RNFL thickness. As in [7], we used the radial basis function as kernel. The SVM was trained by a variation of Platt's sequential minimal optimization algorithm [36]. The SVM hyperparameters were determined by a global optimization technique based on simulated annealing [37] (RNFL-SVM).
2. The ANN classifier of the RNFL thickness. As in [7], we used the multi-layer perceptrons version of the ANN [38] (RNFL-ANN).
3. The proposed method without the MRF *a priori* knowledge on the change detection map (KDS).
4. The proposed method with Gaussian *a priori* knowledge on the free noise ONH image (GBKDS).
5. The proposed change detection method with the fuzzy classifier [9] (fuzzy-DS).
6. The proposed method with the RBF kernel (RBF-BKDS).

It is important to note that we used independent training and test sets to estimate the diagnostic accuracy of the methods. Specifically, we used 10 normal eyes, 5 stable glaucoma eyes for training and 40 normal eyes, 50 stable eyes and 27 glaucoma eyes for testing. The advantage of using stable glaucoma eyes to train the classifier is the fact that the baseline and the follow-up images can be permuted. For example, with one baseline and 4 follow-up images, we can obtain 120 configurations (factorial(5)). This is not the case with normal eyes because the changes due to the normal aging should be integrated in the classification scheme. Because the follow-up duration is relatively short, the changes due to the normal aging are relatively small. Therefore, we used only 10 eyes (20 % of the total normal eyes) to train the classifier. In the case of longer follow-up duration, more normal eyes are required to correctly train the classifier.

Results are presented in Table. 5. The BKDS method with the use of the whole 3D SD-OCT volume instead of the RNFL measurements results in high specificity in both normal and stable glaucoma eyes (94% and 93% respectively) while maintaining good sensitivity (78%). This can be explained by the fact that we take into account the noise for change detection instead of the machine measurements which strongly depend on the image quality. The one class classifier based method led to a superior sensitivity compared to the fuzzy-based classifier [28]. This is likely due to the fact that the behavior of changes due to the glaucoma progression are unknown and hence it is hard to model it. Therefore, it is better to use only the non progression group (e.g; the stable and the normal groups) to train the classifier. Moreover, due in part to the modeling of the correlation among the features which is not taken into account with the basic RBF kernel we obtained a better specificity and sensitivity with the new kernel function.

5. Discussion

Once glaucoma is diagnosed, a sensitive method for detection of progression is essential because appropriate treatment and regular follow-up can slow the disease and preserve vision. The main challenge in the detection of glaucoma progression is to be able to discriminate between glaucomatous structural damage and measurement variability or age-related structural loss. For this reason, it is important to understand the major sources of variability for a given technique and which imaging techniques are the most reproducible. For example, the fundus cameras can be used to estimate the cup to disc ratio. However, in [39], authors showed that the clinically visible disc margin estimated on the fundus images is rarely a consistent anatomic structure and significant errors are present in measurements in patients with glaucoma and healthy controls [40] which limit the use of fundus images for glaucoma progression detection. An alternative to the fundus image is the HRT imaging technique which has been commonly used for glaucoma diagnosis since its commercialization 20 years ago [3]. However, because HRT imaging is limited to ONH surface topography, it cannot measure the RNFL thickness directly [30] which leads to a poor RNFL measurement reproducibility [41]. In [42], authors demonstrated only fair agreement between HRT and clinical judgment of ONH stereophotographs for progression in glaucomatous eyes. At the present time, 3D SD-OCT instruments are now the standard of care for obtaining images of both the ONH and RNFL thickness. Many studies have demonstrated that the retest recognition and eye tracker options in the Spectralis machine

lead to high reproducibility and accurate and repeatable alignment of the baseline and the follow-up images [43–45] which is important for change detection. Therefore, several studies used the SD-OCT peripapillary RNFL thickness measurements (i.e; circular B-scans (3.4-mm diameter, 1536 A-scans) centered at the optic disc were automatically averaged to reduce speckle noise) [46–50] or the ONH measurements (e.g; rim area, minimum rim width, etc.) [51–53] which are either provided by the commercially available spectral-domain optical coherence tomography or manually calculated by experts [52]. In [54], authors used the 3D volume to extract the whole RNFL thickness map instead of the peripapillary RNFL thickness and proposed a variable-size super pixel segmentation to improve the discrimination between early glaucomatous and healthy eyes. In [55], authors used the RNFL thickness map to identify and quantify glaucoma defects.

In this paper, we proposed a new method for detecting glaucoma change using the whole 3D volumes instead of the peripapillary RNFL thickness measurements provided by the instrument. This approach offers several advantages compared with existing methods. First, it allows us to reduce the measurement variability introduced by the noise which is one of the limitations in segmentation of the peripapillary RNFL thickness measurements [8]. Second, the use of the MRF prior to model the spatial dependency favors the generation of homogeneous areas reducing the false positive detection (the change detection method indicated that a change may have occurred when in fact there is no change). Third, the proposed SVDD method allows us to train the classifier using only normal and stable eyes. This is important because we don't have enough knowledge about the pattern of glaucoma change over time. Moreover, training the classifier only on the non-progressing eyes led to a higher specificity. It is important to note that the specificity is generally considered more important in glaucoma progression detection than the sensitivity since the glaucoma is a slow progressing disease [56] and false negatives (people falsely identified as non-progressing) can be eliminated at the next follow-up visits. However, if the specificity is low, this would lead to a substantial follow-up costs, unnecessary treatment (e.g; surgery) and may cause much distress for people who are not progressing. Finally, the proposed algorithm is generic and can be used for different types of eye diseases and imaging modalities.

There are several limitations to the study that should be noted. One limitation of the proposed method is the inclusion of retinal blood vessels in the change detection map. Although the variation of retinal blood vessels may be compensated by the classifier, this compensation by the classification algorithm could cover small changes in the RNFL or rim area. For this reason, some improvement in sensitivity may be accomplished by excluding blood vessels from the change detection map. Another limitation is the lack of a parameter to quantify the rate of progression which is important for clinician's to determine how fast a patient is progressing. Other limitations include the small number of normal and non-progressing eyes to validate the specificity and the sensitivity obtained by the algorithm and the relatively short follow-up time which was on average 2.2 years. Hence, the normal aging effect on the classifier was not evaluated.

6. Conclusion

In this paper, a unified framework for glaucoma progression detection has been proposed. The task of inferring glaucomatous change is tackled with a hierarchical MCMC algorithm that to our knowledge is used for the first time in the glaucoma diagnosis framework. The classification task was formulated as a one class classifier where a hyper-sphere encloses most the unchanged samples (healthy and stable glaucoma eyes). The use of the dependency measurement in the SVDD framework increases the robustness of the proposed change-detection scheme with comparison to the classical SVM and SVDD methods. Of course, any performance gain depends on the quality of the prior samples distribution, which amounts to the quality of chosen distributions and consequently, the copula theory was used. The copula theory provides tools to model samples dependency even if their distribution does not follow a Gaussian distribution such as the SD-OCT images. The validation using clinical data of the proposed approach has shown that it has a high diagnostic accuracy for detecting glaucoma progression compared to other methods.

Acknowledgments

The authors acknowledge the funding and support of the National Eye Institute (grant numbers: U10EY14267, EY019869, EY021818, EY022039 and EY08208, EY11008, P30EY022589 and EY13959).

References

1. Reis, Alexandre SC.; Sharpe, Glen P.; Yang, Hongli; Nicolela, Marcelo T.; Burgoyne, Claude F.; Chauhan, Balwantray C. Optic disc margin anatomy in patients with glaucoma and normal controls with spectral domain optical coherence tomography. *Ophthalmology*. 2012; 119(4):738–747. [PubMed: 22222150]
2. Jaffe GJ, Caprioli J. Optical coherence tomography to detect and manage retinal disease and glaucoma. *American journal of ophthalmology*. 2004; 137(1):156–169. [PubMed: 14700659]
3. Mistlberger A, Liebmann JM, Greenfield DS, Pons ME, Hoh ST, Ishikawa H, et al. Heidelberg retina tomography and optical coherence tomography in normal, ocular-hypertensive, and glaucomatous eyes. *Ophthalmology*. 1999; 106(10):2027–2032. [PubMed: 10519603]
4. Chauhan BC, Blanchard JW, Hamilton DC, LeBlanc RP. Technique for detecting serial topographic changes in the optic disc and peripapillary retina using scanning laser tomography. *Investigative ophthalmology & visual science*. 2000; 41(3):775–782. [PubMed: 10711693]
5. Balasubramanian M, Kriegman DJ, Bowd C, Holst M, Weinreb RN, Sample PA, et al. Localized glaucomatous change detection within the proper orthogonal decomposition framework. *Investigative Ophthalmology & Visual Science*. 2012; 53(7):3615–3628. [PubMed: 22491406]
6. Belghith, Akram; Bowd, Christopher; Balasubramanian, Madhusudhanan; Weinreb, Robert N.; Zangwill, Linda M. A bayesian framework for glaucoma progression detection using heidelberg retina tomograph images. *International Journal of Advanced Computer Science & Applications*. 2013; 4(9):223–229.
7. Bizios D, Heijl A, Hougaard J Leth, Bengtsson B. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by stratus oct. *Acta ophthalmologica*. 2010; 88(1):44–52. [PubMed: 20064122]
8. Balasubramanian M, Bowd C, Vizzeri G, Weinreb RN, Zangwill LM. Effect of image quality on tissue thickness measurements obtained with spectral-domain optical coherence tomography. *Optics express*. 2009; 17(5):4019–4036. [PubMed: 19259243]
9. Belghith, Akram; Bowd, Christopher; Weinreb, Robert N.; Zangwill, Linda M. SPIE Medical Imaging. International Society for Optics and Photonics; 2014. A joint estimation detection of glaucoma progression in 3d spectral domain optical coherence tomography optic nerve head images; p. 903-906.

10. Li, SZ. Markov random field modeling in computer vision. Springer-Verlag New York, Inc; 1995.
11. McLachlan, GJ.; Krishnan, T. The EM algorithm and extensions. Vol. 274. Wiley; New York: 1997.
12. Gilks, Walter R.; Richardson, Sylvia; Spiegelhalter, David J. Markov Chain Monte Carlo in Practice. Chapman and Hall; London: 1996.
13. Shawe-Taylor, J.; Cristianini, N. Kernel methods for pattern analysis. Cambridge University Press; Cambridge: 2004.
14. Tax DMJ, Duin RPW. Support vector data description. Machine Learning. 2004; 54(1):45–66.
15. Enzweiler M, Gavrilu DM. Monocular pedestrian detection: Survey and experiments. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009; 31(12):2179–2195. [PubMed: 19834140]
16. Camps-Valls G, Gómez-Chova L, Muñoz-Marí J, Rojo-Álvarez JL, Martínez-Ramón M. Kernel-based frame work for multitemporal and multisource remote sensing data classification and change detection. IEEE Transactions on Geoscience and Remote Sensing. 2008; 46(6):1822–1835.
17. Belghith A, Collet C, Armspach JP. Change detection based on a support vector data description that treats dependency. Pattern Recognition Letters. 2012; 34(3):275–282.
18. Claeskens, Gerda; Hjort, Nils Lid. Model selection and model averaging. Cambridge University Press; Cambridge: 2008.
19. Jeffrey, A.; Zwillinger, D. Table of integrals, series, and products. Academic Press; Waltham, Massachusetts: 2007.
20. Robert, Christian. Discretization and MCMC convergence assessment. Vol. 135. Springer-Verlag; New York: 1998.
21. Gelman A, Roberts G, Gilks W. Efficient metropolis jumping rules. Bayesian statistics. 1996; 5:599–608.
22. Gilks, Walter R.; Roberts, Gareth O.; Sahu, Sujit K. Adaptive markov chain monte carlo through regeneration. Journal of the American Statistical Association. 1998; 93(443):1045–1054.
23. Joe, H. Multivariate models and dependence concepts. Chapman & Hall; London: 1997.
24. Johansen, Soren; Juselius, Katarina. Maximum likelihood estimation and inference on cointegration-with applications to the demand for money. Oxford Bulletin of Economics and statistics. 1990; 52(2):169–210.
25. Cawley GC, Talbot NLC. Efficient leave-one-out cross-validation of kernel Fisher discriminant classifiers. Pattern Recognition. 2003; 36(11):2585–2592.
26. Na, Jung Hwa; Sung, Kyung Rim; Baek, Seunghee; Lee, Jin Young; Kim, Soa. Progression of retinal nerve fiber layer thinning in glaucoma assessed by cirrus optical coherence tomography-guided progression analysis. Current eye research. 2013; 38(3):386–395. [PubMed: 23441595]
27. Gardiner, Stuart K.; Ren, Ruojin; Yang, Hongli; Fortune, Brad; Burgoyne, Claude F.; Demirel, Shaban. A method to estimate the amount of neuroretinal rim tissue in glaucoma: Comparison with current methods for measuring rim area. American journal of ophthalmology. 2014; 157(3): 540–549. [PubMed: 24239775]
28. Belghith, Akram; Bowd, Christopher; Weinreb, Robert N.; Zangwill, Linda M. A hierarchical framework for estimating neuroretinal rim area using 3d spectral domain optical coherence tomography (sd-oct) optic nerve head (onh) images of healthy and glaucoma eyes,” in. Engineering in Medicine and Biology Society. 2014:3869–3872.
29. Wang, Xun; Fan, Yinghao; Kong, Dingke. A level set sar image segmentation approach based on fisher distribution. Journal of Convergence Information Technology (JCIT). 7(21):252–2012.
30. Volpi, Michele; Tuia, Devis; Camps-Valls, Gustavo; Kanevski, Mikhail. Unsupervised change detection with kernels. Geoscience and Remote Sensing Letters, IEEE. 2012; 9(6):1026–1030.
31. Bazi Y, Bruzzone L, Melgani F. An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images. IEEE Transactions on Geoscience and Remote Sensing. 2005; 43(4):874–887.
32. Cohen, Jacob. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960; 20:37–46.

33. Brennan, Robert L.; Prediger, Dale J. Coefficient kappa: Some uses, misuses, and alternatives. Educational and psychological measurement. 1981; 41(3):687–699.
34. Masson, Pascale; Pieczynski, Wojciech. Sem algorithm and unsupervised statistical segmentation of satellite images. Geoscience and Remote Sensing, IEEE Transactions on. 1993; 31(3):618–633.
35. Besag J. On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society. Series B (Methodological). 1986; 48(3):259–302.
36. Fan RE, Chen PH, Lin CJ. Working set selection using second order information for training support vector machines. The Journal of Machine Learning Research. 2005; 6:1889–1918.
37. Imbault, F.; Lebart, K. A stochastic optimization approach for parameter tuning of support vector machines. Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. IEEE; 2004; p. 597–600.
38. Gardner MW, Dorling SR. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric environment. 1998; 32(14–15):2627–2636.
39. Reis, Alexandre SC.; Sharpe, Glen P.; Yang, Hongli; Nicolela, Marcelo T.; Burgoyne, Claude F.; Chauhan, Balwantray C. Optic disc margin anatomy in patients with glaucoma and normal controls with spectral domain optical coherence tomography. Ophthalmology. 2012; 119(4):738–747. [PubMed: 22222150]
40. Reis, Alexandre SC.; O’Leary, Neil; Yang, Hongli; Sharpe, Glen P.; Nicolela, Marcelo T.; Burgoyne, Claude F., et al. Influence of clinically invisible, but optical coherence tomography detected, optic disc margin anatomy on neuroretinal rim evaluation. Investigative ophthalmology & visual science. 2012; 53(4):1852–1860. [PubMed: 22410561]
41. Miglior, Stefano; Albé, Elena; Guareschi, Magda; Rossetti, Luca; Orzalesi, Nicola. Intraobserver and interobserver reproducibility in the evaluation of optic disc stereometric parameters by heidelberg retina tomograph. Ophthalmology. 2002; 109(6):1072–1077. [PubMed: 12045046]
42. Kourkoutas, Dimitrios; Buys, Yvonne M.; Flanagan, John G.; Hatch, Wendy V.; Balian, Carmen; Trope, Graham E. Comparison of glaucoma progression evaluated with heidelberg retina tomograph ii versus optic nerve head stereophotographs. Canadian Journal of Ophthalmology/ Journal Canadien d’Ophtalmologie. 2007; 42(1):82–88.
43. Leary, Neil OO.; Crabb, David P.; Mansberger, Steven L.; Fortune, Brad; Twa, Michael D.; Lloyd, Michael, et al. Glaucomatous progression in series of stereoscopic photographs and heidelberg retina tomograph images. Archives of ophthalmology. 2010; 128(5):560–568. [PubMed: 20457976]
44. Langenegger, Stefan J.; Funk, Jens; Töteberg-Harms, Marc. Reproducibility of retinal nerve fiber layer thickness measurements using the eye tracker and the retest function of spectralis sd-oct in glaucomatous and healthy control eyes. Investigative ophthalmology & visual science. 2011; 52(6):3338–3344. [PubMed: 21330656]
45. Pemp, Berthold; Kardon, Randy H.; Kircher, Karl; Pernicka, Elisabeth; Schmidt-Erfurth, Ursula; Reitner, Andreas. Effectiveness of averaging strategies to reduce variance in retinal nerve fibre layer thickness measurements using spectral-domain optical coherence tomography. Graefes’s Archive for Clinical and Experimental Ophthalmology. 2013; 251(7):1841–1848.
46. Mwanza, Jean-Claude; Oakley, Jonathan D.; Budenz, Donald L.; Anderson, Douglas R. Ability of cirrus hd-oct optic nerve head parameters to discriminate normal from glaucomatous eyes. Ophthalmology. 2011; 118(2):241–248. [PubMed: 20920824]
47. Cho, Jung Woo; Sung, Kyung Rim; Hong, Jung Taeck; Um, Tae Woong; Kang, Sung Yong; Kook, Michael S. Detection of glaucoma by spectral domain-scanning laser ophthalmoscopy/optical coherence tomography (sd-slo/oct) and time domain optical coherence tomography. Journal of glaucoma. 2011; 20(1):15–20. [PubMed: 20436370]
48. Seong, Mincheol; Sung, Kyung Rim; Choi, Eun Hee; Kang, Sung Yong; Cho, Jung Woo; Um, Tae Woong; Kim, Yoon Jeon; Park, Seong Bae; Hong, Hun Eui; Kook, Michael S. Macular and peripapillary retinal nerve fiber layer measurements by spectral domain optical coherence tomography in normal-tension glaucoma. Investigative ophthalmology & visual science. 2010; 51(3):1446–1452. [PubMed: 19834029]

49. Leite, Mauro T.; Rao, Harsha L.; Zangwill, Linda M.; Weinreb, Robert N.; Medeiros, Felipe A. Comparison of the diagnostic accuracies of the spectralis, cirrus, and rtvue optical coherence tomography devices in glaucoma. *Ophthalmology*. 2011; 118(7):1334–1339. [PubMed: 21377735]
50. Kotowski, Jacek; Folio, Lindsey S.; Wollstein, Gadi; Ishikawa, Hiroshi; Ling, Yun; Bilonick, Richard A.; Kagemann, Larry; Schuman, Joel S. Glaucoma discrimination of segmented cirrus spectral domain optical coherence tomography (sd-oct) macular scans. *British Journal of Ophthalmology*. 2012; 96(11):1420–1425. [PubMed: 22914498]
51. Gardiner, Stuart K.; Ren, Ruojin; Yang, Hongli; Fortune, Brad; Burgoyne, Claude F.; Demirel, Shaban. A method to estimate the amount of neuroretinal rim tissue in glaucoma: Comparison with current methods for measuring rim area. *American journal of ophthalmology*. 2014; 157(3): 540–549. [PubMed: 24239775]
52. Chauhan, Balwantray C.; O'Leary, Neil; AlMobarak, Faisal A.; Reis, Alexandre SC.; Yang, Hongli; Sharpe, Glen P., et al. Enhanced detection of open-angle glaucoma with an anatomically accurate optical coherence tomography-derived neuroretinal rim parameter. *Ophthalmology*. 2013; 120(3):535–543. [PubMed: 23265804]
53. Patel, Nimesh; Sullivan-Mee, Michael; Harwerth, Ronald S. The relationship between retinal nerve fiber layer thickness and optic nerve head neuroretinal rim tissue in glaucoma. *Investigative ophthalmology & visual science*. 2014; 55(10):6802–6816. [PubMed: 25249610]
54. Xu, Juan; Ishikawa, Hiroshi; Wollstein, Gadi; Bilonick, Richard A.; Folio, Lindsey S.; Nadler, Zach, et al. Three-dimensional spectral-domain optical coherence tomography data analysis for glaucoma detection. *PloS one*. 2013; 8(2):e55476. [PubMed: 23408988]
55. Shin, Joong Won; Uhm, Ki Bang; Seong, Mincheol; Kim, Yu Jeong. Diffuse retinal nerve fiber layer defects identification and quantification in thickness maps. *Investigative ophthalmology & visual science*. 2014; 55(5):3208–3218. [PubMed: 24744205]
56. Weinreb, Robert N.; Khaw, Peng Tee. Primary open-angle glaucoma. *The Lancet*. 2004; 363(9422):1711–1720.

Highlights (for review)

- We present a two-step framework for glaucoma progression detection
- The Markov prior increases significantly the change detection rate
- The one-class kernel classifier performs better than the fuzzy-based classifier
- The use of the whole 3D SD-OCT images improves the glaucoma change detection

Figure 1.
Comparison of the goodness of fit of gamma, log-normal, Weibull and normal distributions for retinal intensities.

Figure 2.
The RNFL and the rim area in a B-scan.

Figure 3.
The hypersphere enclosing most of the targets estimated using the SVDD approach.

Figure 4.
The hyperplane separating most of the targets from the outliers estimated using the SVM approach.

Figure 5. Example of a simulated image: (a) simulated image with PSNR=20 dB and (b) the corresponding ground truth: the white boxes represent the changes.

Figure 6.

Example of semi-simulated image: (a) the original B-scan image (b) image semi-simulated by the permutation of image regions (c) the ground truth: the white boxes represent the changes.

Table 1

Description of the clinical dataset used to assess the diagnosis accuracy of the proposed framework.

	Number of patients	Number of eyes	Follow-up (median/interquartile range)
Normal group	23	40	2.03 (1.8) years
Stable glaucoma group	50	26	5 (0) weeks
Progressing glaucoma group	27	27	2.4 (1.6) years

Table 2

Description of the simulated and semi-simulated dataset used to evaluate the proposed change detection method.

	Number of images
Simulated dataset	100
Semi-simulated dataset by region permutation	100
Semi-simulated dataset by intensity modification	100

Table 3

Mean and standard deviation (SD) of false detection, missed detection, total errors and kappa coefficient resulting from: the proposed Bayesian Markovian approach, the kernel k-means method and the threshold method using simulated and semi-simulated.

Semi-simulated datasets	False detection (±SD) %	Missed detection (±SD) %	Total errors (±SD) %	Kappa
Bayesian Markovian approach	0.54 (±0.21) %	5.62 (±1.03) %	0.91 (±1.03) %	78
Kernel k-means	0.89 (±0.41) %	8.03 (±2.75) %	2.03 (±0.51) %	71**
Threshold	2.14 (±0.71) %	10.52 (±3.14) %	3.45 (±1.24) %	64**

Simulated dataset	False detection (±SD) %	Missed detection (±SD) %	Total errors (±SD) %	Kappa
Bayesian Markovian approach	1.24 (±0.35) %	7.05 (±1.95) %	1.51 (±0.34) %	74
Kernel k-means	1.65 (±0.63) %	8.89 (±2.87) %	2.38 (±2.44) %	67**
Threshold	3.78 (±1.47) %	15.25 (±SD4.17) %	6.41 (±2.74) %	63**

** indicates that the change detection results are statistically different from the Bayesian Markovian approach (p-value< 0.05 using paired t-test).

Table 4

Mean and SD false detection, missed detection, total errors and the kappa coefficient resulting from: the proposed MCMC procedure, the SEM procedure and the ICM procedure using the simulated dataset.

Simulated dataset	False detection ($\pm$ SD) %	Missed detection ($\pm$ SD) %	Total errors ($\pm$ SD) %	Kappa
Proposed MCMC	1.24 ($\pm$ 0.35) %	7.05 ($\pm$ 1.95) %	1.51 ($\pm$ 0.34) %	70
SEM	1.74 ($\pm$ 0.57) %	7.94 ($\pm$ 2.14) %	1.82 ($\pm$ 0.39) %	64**
ICM	1.85 ($\pm$ 0.47) %	8.24 ($\pm$ 2.10) %	1.95 ($\pm$ 0.44) %	63**

** indicates that the change detection results are statistically different from the Bayesian Markovian approach (p-value< 0.05 using paired t-test).

Table 5

Diagnostic accuracy of different methods.

	Progressing glaucoma group sensitivity	Healthy group specificity	Stable glaucoma group specificity	AUROC
BKDS	78 %	93 %	94 %	0.91
KDS	70 %	84 %	81 %	0.79
Fuzzy-DS	64 %	90 %	92 %	0.83
GBKDS	52 %	71 %	78 %	0.65
RBF-BKDS	72 %	83 %	80 %	0.76
RNFL-ANN	69 %	71 %	78 %	0.69
RNFL-SVM	52 %	68 %	79 %	0.60