Locally Adaptive Probabilistic Models for Global Segmentation of Pathological OCT Scans

Fabian Rathke[†2], Mattia Dessana[†1] and Christoph Schnörr[1,2]

[1] University of Heidelberg, Image & Pattern Analysis Group (IPA)
[2] Heidelberg Collaboratory for Image Processing (HCI)

Abstract. Segmenting retinal tissue deformed by pathologies can be challenging. Segmentation approaches are often constructed with a certain pathology in mind and may require a large set of labeled pathological scans, and therefore are tailored to that particular pathology.
We present an approach that can be easily transfered to new pathologies, as it is designed with no particular pathology in mind and requires no pathological ground truth. The approach is based on a graphical model trained for healthy scans, which is modified *locally* by adding pathology-specific shape modifications. We use the framework of sum-product networks (SPN) to find the best combination of modified and unmodified local models that *globally* yield the best segmentation. The approach further allows to localize and quantify the pathology. We demonstrate the flexibility and the robustness of our approach, by presenting results for three different pathologies: diabetic macular edema (DME), age-related macular degeneration (AMD) and non-proliferative diabetic retinopathy.

1   Introduction

Since its introduction in 1991, Optical Coherence Tomography has established itself as an invaluable diagnostic tool. Early publications focused on the segmentation of healthy retina scans. Various approaches have been devised, rendering this problem more or less solved. Consequently, recent publications shifted their attention to the pathological case, with the aim of detecting, assessing and monitoring theses diseases as accurately as possible.

Related work. AMD, a much addressed pathology, e.g. [1], is characterized by cellular debris called drusen, that accumulates between the retinal pigment epithelium (RPE) and the underlying choroid, with size and number indicating the stage of the disease (example in Fig. 1 (b)). While leading to considerable deformations of the retina, cell layers themselves stay more or less intact in the early and intermediate stages, considered in the segmentation literature.

Two recent works addressed the more involved problem of segmenting individual layers in the presence of DME [2, 3]. Here fluid leaking from damaged

---

† Both authors contributed equally.

(a) Mild Diabetic RP   (b) Intermediate AMD   (c) Advanced DME
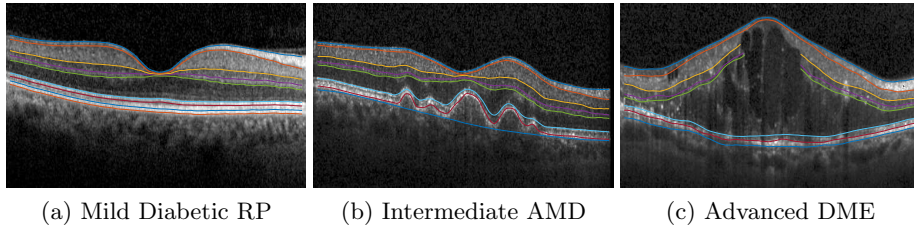
Fig. 1: Three pathologies of different difficulty all segmented by the same model, by adding pathology-specific shape information to a model for healthy scans.

capillaries leads to a swelling of the macular, possibly leading to the destruction of some layers in the retina (Fig. 1 (c)). The dataset used in both publications above is very challenging in that regard, and constitutes a harder problem than intermediate AMD. Both approaches rely on ground truth to a) train a classifier to find fluid regions and exclude them from the segmentation [2] or b) train DME specific appearance models [3].

Contribution. We present an approach that is adaptable to different pathologies *without the need for ground truth.* This is done by using a locally adaptive extended version of a graphical model trained on healthy data. Based on a sum-product network that enables tractable globally optimal inference, we find the optimal combination of regions and modifications and combine them into a final segmentation. We demonstrate its capabilities by segmenting three different pathologies. Besides obtaining accurate segmentations, we are also able to localize the pathological regions. To our knowledge, this is the first time, that a single approach was tested and evaluated for more than one pathology. The basis of our approach is sketched in Section 2, followed by introducing our approach in Section 3 and discussing results in Section 4.

2   Probabilistic Graphical Model

Adopting our previous approach [4], we model an OCT scan $y \in \mathbb{R}^{M \times N}$ (N A-Scans with M pixels each) and its segmentations b and c respectively. Here $c \in \mathbb{N}^{K \cdot N}$ (K boundaries) denotes the discretized version of the continuous boundary vector $b \in \mathbb{R}^{K \cdot N}$, which is the connection between the discrete pixel domain of y and the continuous boundary domain of b. The graphical model is given by

$$p(y, c, b) = p(y|c)p(c|b)p(b). \tag{1}$$

We will briefly discuss each component, and refer to [4] for more details.

Appearance $p(y|c)$. Appearance of boundaries and layers is modeled via *local* class-specific Gaussian densities: The probability of pixel $y_{i,j}$ belonging to a class
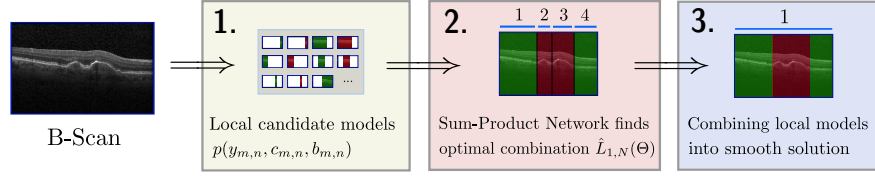
Fig. 2: Workflow: [1.] We segment many *local* graphical models (introduced in Sec. 2) being either modified (red) or unmodified (green) (see (8)) for various subregions of the B-Scan. [2.] We then find the *globally* optimal combination $\hat{L}_{1,n}(\Theta)$ (10) using SPNs (Sec. 3, first part). [3.] Finally, we fuse the local models of $\hat{L}_{1,n}(\Theta)$ into a smooth segmentation (Sec. 3, second part).

$x_{i,j} \in \{l_1, \ldots, l_n, t_1, \ldots, t_{n-1}\}$ (see Fig. 3 (a)) is modeled as Gaussian,

$$p(y|c) = \prod_{i=1}^{M} \prod_{j=1}^{N} p(y_{i,j}|c), \qquad p(y_{i,j}|c) = \mathcal{N}(\tilde{y}_{i,j}|\mu_{x_{i,j}}, \Sigma_{x_{i,j}}), \qquad (2)$$

where the class-label $x_{i,j}$ is determined by the boundary configuration c and $\tilde{y}_{i,j}$ is a patch around pixel $y_{i,j}$.

Shape $p(b)$. The *global* shape prior captures typical variations of cell layer boundaries. The shape vector b is determined by a linear Gaussian model

$$b = Ws + \mu + \epsilon, \quad s \sim \mathcal{N}(0, I), \quad \epsilon \sim \mathcal{N}(0, \sigma^2). \qquad (3)$$

The matrix $W \in \mathbb{R}^{K \cdot N \times m}$ maps the low-dimensional vector $s \in \mathbb{R}^m$ onto b. Each column of W denotes a certain shape variation that gets added to the mean shape μ. Given n training segmentations $X \in \mathrm{R}^{n \times N \cdot K}$, W is obtained by the first m eigenvectors of $\text{cov}(X)$ weighted by the corresponding eigenvectors, and μ simply is $\overline{X}$. The marginal distribution of b can then be shown to be

$$p(b) = \mathcal{N}(b; \mu, \Sigma = WW^T + \sigma^2 I). \qquad (4)$$

MRF Regularization $p(c|b)$. Shape and appearance interact in a Markov random field over the discrete variable c. It is composed of column-wise chain models that allow for parallel inference (with more details to be found in [4])

$$p(c|b) = \prod_{j=1}^{N} p(c_j|b), \qquad p(c_j|b) = p(c_{1,j}|b) \prod_{k=2}^{K} p(c_{k,j}|c_{k-1,j}, b). \qquad (5)$$

Inference. In [4] we proposed a variational scheme: Design a tractable graphical model $q(c, b)$ by adding conditional independences, then infer the full distribution $q(b, c)$ by minimizing the Kullback-Leibler (KL) divergence to $p(c, b|y)$. We decoupled the discrete and continuous model components, $q(c, b) = q_c(c)q_b(b)$ while keeping the remaining structure intact: That is $q_c(c)$ are column-wise MRFs as

in (5) and $q(b) = \mathcal{N}(b; \bar{\mu}, \bar{\Sigma})$. Infering $q(c, b)$ then corresponds to minimizing the following non-convex optimization problem

$$\min_{q_c, \bar{\mu}, \bar{\Sigma}} \mathrm{KL}(q(c,b) \| p(c,b|y)) = \int_b \sum_c q(c,b) \log \frac{q(c,b)}{p(c,b|y)}. \tag{6}$$

Plugging in the definitions of $q(c, b)$ and $p(c, b|y)$, one can find explicit update equations for the parameters of q_c and q_b. Of interest for this work is the update step for $\bar{\mu}$, which is of the form

$$A(\bar{\mu} - \mu) = \mathbb{E}_{q_c}[c] - \mu \quad \Longrightarrow \quad A\bar{\mu} = (A - I)\mu + \mathbb{E}_{q_c}[c]. \tag{7}$$

It links the mean of q_c ($\mathbb{E}_{q_c}[c]$) to the mean of q_b ($\bar{\mu}$) via the linear mapping A determined from Σ. We will revert to this equation at the end of the next Section, see (12). The optimization alternates between solving the MRF q_c and updating the parameters of q_b until convergence.

3 Locally Adaptive Priors

The model described above, when trained on healthy data, is not sufficiently flexible to adapt to unseen pathologies with large deformations. We address this problem by finding a global optimal combination of locally modified submodels using the principle of maximum-likelihood and dynamic programming.

Sum-Product Networks. We assume that models of pathological structure are *translation invariant*, *local* and *approximately independent*. Independence and locality allow to factorize the full distribution $p(y, b, c)$ into local distributions, an assumption necessary for SPNs. Translation invariance implies that the pathology can appear at any horizontal position in the image.

Recall that W in (3) contains typical shape variations of healthy retina layers. We adapt the graphical model towards an illness, by adding translation-invariant pathology-specific modes W^{ill} to W:

$$\theta^{\mathrm{ill}}_{m,n} := \left(W_{m,n}\ W^{\mathrm{ill}}_{m,n} \right), \qquad \theta^{\mathrm{healthy}}_{m,n} := W_{m,n}. \tag{8}$$

Here subscript m, n denote the pruning of W to the region $[m, n]$. Segmenting various such models for different regions $[m, n]$ with ill and healthy parameters constitutes step 1 in our workflow (Fig. 2).

Let $L_{m,n}(\theta^z_{m,n})$ be the log-likelihood of the segmentation for region $[m, n]$:

$$L_{m,n}(\theta^z_{m,n}) := \log q \left(c_{m,n}, b_{m,n} | y_{m,n}, \theta^z_{m,n} \right), \qquad z \in \{\text{healthy}, \text{ill}\}. \tag{9}$$

Now let $X = \{1, x_1, x_2, ..., x_H, N\}$ denote the division of y into $H + 1$ regions and let $\Theta = \left\{ \theta^{z_1}_{1,x_1}, \theta^{z_2}_{x_1,x_2}, ..., \theta^{z_{H+1}}_{x_H,N} \right\}$ denote the corresponding set of shape modifications. We want to find the combination of submodels with maximal total log-likelihood

$$\hat{L}_{1,N}(\Theta) = \underset{H, X, \theta}{\operatorname{argmax}}\ L_{1,x_1}(\theta^{z_1}_{x_1, x_2}) + ... + L_{x_H, N}(\theta^{z_{H+1}}_{x_H, N}). \tag{10}$$

The global optimum of this combinatorial problem can be found with *dynamic programming*. Let $\hat{L}_{m,n}$ be the optimal selection of X and Θ in region $[m, n]$. It can be computed recursively as:

$$\hat{L}_{m,n} = \max \left(\max_{x \in \{m,m+1,\dots,n\}} \left(\hat{L}_{m,x} + \hat{L}_{x,n} \right), \max_{z \in \{\text{ill,healthy}\}} L_{m,n}(\theta_{m,n}^z) \right), \quad (11)$$

which is the maximum between the single best model over region $[m, n]$ and the optimal factorization in two adjacent areas. To compute $\hat{L}_{m,n}$ for regions of width w, we need quantities $\hat{L}_{m,x}, \hat{L}_{x,n}$ for all regions of width $< w$. Given $\hat{L}_{m,x}$ and $\hat{L}_{x,n}$, the complexity is dominated by evaluating $L_{m,n}(\theta_{m,n}^z)$ (9).

Assuming a minimal width w_{min}, this suggests an iterative algorithm: first, compute $\hat{L}_{m,n}$ for regions of width w_{min}. Then, recursively compute (11) for regions of increasingly higher w. We can reduce complexity even further, by increasing and shifting windows with some fixed step size $s > 1$. Due to the nature of dynamic programming, many terms $\hat{L}_{m,n}$ get reused during the optimization. To favor more compact subregions, we add a regularization to (10) to punish models of small size. This algorithm implements globally optimal MAP-inference in a SPN [5] and constitutes step 2 in our workflow (Fig. 2).

Combining Local Models. Because submodels are found independently, they usually constitute a non-smooth segmentation of y. To obtain a smooth solution, we solve a modified version of the full graphical model $p(y, b, c)$ taking into account the optimal solution (10), corresponding to step 3 in Fig. 2.

The MAP estimate (10) can be interpreted as a graphical model $p(y, c, b)$ without coupling between subregions. This can be enforced by setting all entries in Σ to zero that belong to boundary positions in two different regions. Furthermore, for subregions identified as ill, we use the modified shape modes $\theta_{m,n}^{\text{ill}}$ to calculate the submatrix of Σ via (4). Solving the full graphical model with such a modified covariance matrix would yield the same segmentation as $\hat{L}_{1,N}(\Theta)$.

Now to enforce smoothing while staying close to the SPN solution, we replace the system of linear equations (7) by the constrained least-squares problem:

$$\min_{\bar{\mu}} \| \tilde{A}\bar{\mu} - (A - I)\mu - \mathbb{E}_{q_c}[c] \|^2, \qquad \text{subject to} \quad B\bar{\mu} \leq \delta \mathbb{1}, \qquad (12)$$

where $\mathbb{1}$ is a vector of ones. Each row in the constraint matrix B selects two neighboring entries in $\bar{\mu}$ belonging to two different subregions and restricts their difference to be less than δ. This enforces a weak coupling between subregions. Solving the full graphical model with the sparse Σ and the modified update step for $\bar{\mu}$ then yields a smooth segmentation, as Fig. 3 (b) and (c) demonstrates.

4 Results

We demonstrate the flexibility of our approach by segmenting three different pathologies, ranging from minor deformations to severe distortions of the retina structure. We will use the same graphical model for all pathologies, only adapting
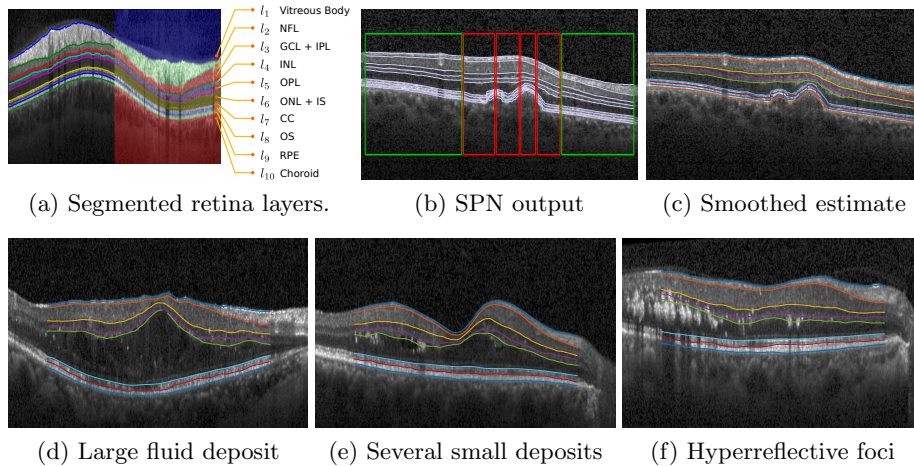
(a) Segmented retina layers. (b) SPN output (c) Smoothed estimate

(d) Large fluid deposit (e) Several small deposits (f) Hyperreflective foci

Fig. 3: (a) The names of the segmented retina layers. Surfaces 1-9 lie in between layers $l_1, \ldots, l_{10}$. (b) A SPN estimate and its smoothed version (c). Note that Bruch's membrane (surface 9) gets fitted in a post-processing step, described in the results section. (d)-(f) Example segmentations from the DME dataset.

the pathological shape modes we add. During inference the graphical model can pick the strength and the sign of any mode freely.

We trained the healthy model on the same 35 labeled volumes also used in [4]. As prediction we used the expectation $\mathbb{E}_{q_c}[c]$ of c. The error metric is the unsigned error between labels and the prediction, averaged over all B-Scans in a volume and A-Scans therein. We used $\delta = 2$ pixel (see (12)) throughout our experiments. Table 1 summarizes all results.

Diabetic Retinopathy. The dataset of [6] contains 10 subjects (5 B-Scans each) affected by mild non-proliferative diabetic retinopathy (RP). As only small deformations occur, we used our graphical model of [4]. Since the dataset lacked relative positions of B-Scans inside the volume, which we require to select a shape prior, we estimated the position as following. For each B-Scan we tested all shape priors and a) used the one with the largest model likelihood and b) the one with the smallest error. This yielded a lower and upper bound on the true error, if the information would have been available, which we averaged for the final result.

AMD. We used an in-house dataset with 8 Spectralis volumes of early and intermediate AMD and labels for surfaces 1, 8 and 9 for all 19 B-Scans. We added one mode with the sine function evaluated between 0 and π for surfaces 6-9, simulating the effect of those layers being pushed up by a circular-shaped fluid deposit underneath. While Bruch's membrane (surface 9) is supposed to lie beneath the fluid region, better segmentations where obtained if it was included in the shape mode. The final segmentation for this surface was given by the

Table 1: Unsigned error for all tested datasets in μm (1px = 3.87 μm). Surface numbers 1–9 correspond to Fig. 3 (a). '–' marks the absence of labels.

Dataset	Method	Avg.	1	2	3	4	5	6	8	9
RP	Tian et al. [6]	4.48	**3.70**	4.49	3.84	–	5.75	–	–	4.63
	Our method	**4.08**	4.39	**4.15**	3.84	–	**4.65**	–	–	**3.37**
AMD	Our method	4.90	2.87	–	–	–	–	–	6.06	5.77
DME 1-5	Chiu et al. [2]	7.82	6.59	8.38	9.04	**11.02**	**11.01**	4.84	5.74	5.91
	Karri et al. [3]	9.54	**4.47**	11.77	11.12	17.54	16.74	4.99	**5.35**	**4.30**
	Our method	**7.71**	4.66	**6.78**	**8.87**	11.02	13.60	**4.61**	7.06	5.11
DME 6-10	Chiu et al. [2]	5.81	5.01	6.37	7.46	7.34	7.74	3.88	4.34	4.32
	Karri et al. [3]	5.14	3.64	5.95	6.48	**6.64**	8.00	**3.09**	**4.12**	3.17
	Our method	**5.11**	**3.62**	**4.87**	**5.92**	7.50	**7.69**	3.29	4.83	**3.16**

conditional mean $\mu_{a|b} = \mu_a - (K_{aa})^{-1} K_{ab}(x_b - \mu_b)$ of (4), where x_b denotes the part of the segmentation identified as healthy.

DME. The dataset published by Chiu et al. [2] consists of 10 Spectralis volumes with 11 labeled B-Scans per volume. While volumes 6-10 are mild and intermediate cases, volumes 1-5 constitute advanced DME cases, with disappearing layers (Fig. 1 (c)) and advanced texture artifacts due to highly reflective regions characteristic for DME (Fig. 3 (c)).

To reduce sensitivity to the texture artifacts, we added patches of size 7×7 and 3×3 (besides the standard 15×15 patches). To deal with the disappearing layers, we dropped the segmentation in regions of low intensity if the difference between surface 1 and 9 exceeded a threshold. As pathology-specific modes we added a set of connected linear functions to boundaries 1-5, which could only be adjusted jointly. Furthermore, as DME can be accompanied by a swelling of the nerve fiber layer (NFL), we added linear functions to surfaces 1 and 2.

Karri et al. [3] also tested their approach on this dataset, but only published results for volumes 6-10, using the first 5 volumes for training. Using their published code (`https://github.com/ultrai/Chap_1`), we could reproduce their results for volumes 6-10, as well as reverse training and test set to obtain results for volumes 1-5. Results are the displayed in the lower half of Table 1. For a fair comparison, we also applied the mechanism for dropping segmentations.

In general, less difficult volumes 6-10 yield lower errors for all approaches as expected. Karri's and our approach perform best. The situation changes for the more difficult volumes 1-5. Now Chiu's and our approach perform on par, beating the one of Karri et al., which lacks sufficient shape regularization [3].

Pathology hinting. Fig. 4 demonstrates another benefit of using a shape prior. Given a segmentation b, one can calculate the latent variable s, which indicates how much each mode was utilized (3). The red surfaces indicate the usage of pathological modes $W_{m,n}^{ill}$, plotted below the lowest boundary affected.
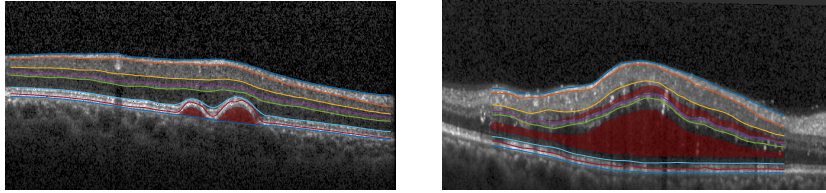
Fig. 4: Estimates of fluid regions due to the pathological modes $W_{m,n}^{\mathrm{ill}}$ used.

5   Discussion

We presented a method for the segmentation of pathological OCT data, combining a graphical model and sum-product networks. While our approach yields state-of-the-art performance, it does not require labeled ground truth data. Furthermore, it can segment *several* pathologies. To our knowledge, this is a feature not demonstrated yet by any other approach. Last but not least, it can *localize* the pathological area, which could be valuable for practitioners. An evaluation of this feature will be part of our future work

The current approach was evaluated in 2-D, requiring between 30 and 60 seconds per B-Scan. While all parts of our workflow naturally extend to 3-D, the number of submodels in step 1 grows exponentially, making a direct conversion too costly. Future work may include mechanisms to prune the SPN search, reducing the amount of tested submodels. This would benefit the current 2-D as well as any potential 3-D approach.

References

1. Chiu, S.J., Izatt, J.A., O'Connell, R.V., Winter, K.P., Toth, C.A., Farsiu, S.: Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. Invest. Ophthalmol. Vis. Sci. **53**(1) (2012) 53
2. Chiu, S.J., Allingham, M.J., Mettu, P.S., Cousins, S.W., Izatt, J.A., Farsiu, S.: Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema. Biomed. Opt. Express **6**(4) (2015) 1172–1194
3. Karri, S., Chakraborthi, D., Chatterjee, J.: Learning layer-specific edges for segmenting retinal layers with large deformations. Biomed. Opt. Express **7**(7) (2016) 2888–2901
4. Rathke, F., Schmidt, S., Schnörr, C.: Probabilistic intra-retinal layer segmentation in 3-D OCT images using global shape regularization. Med. Image Anal. **18**(5) (2014) 781–794
5. Poon, H., Domingos, P.: Sum-product networks: A new deep architecture. In: UAI. (2011) 337–346
6. Tian, J., Varga, B., Tatrai, E., Fanni, P., Somfai, G.M., Smiddy, W.E., Debuc, D.C.: Performance evaluation of automated segmentation software on optical coherence tomography volume data. J. Biophotonics **9**(5) (2016) 478–489